

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
4 August 2005 (04.08.2005)

PCT

(10) International Publication Number
WO 2005/069900 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2005/001469

(22) International Filing Date: 14 January 2005 (14.01.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/537,341 16 January 2004 (16.01.2004) US

(71) Applicant (for all designated States except US): MERCK
& CO., INC. [US/US]; 126 East Lincoln Avenue, RY
60-30, Rahway, NJ 07065 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): GARCIA-CALVO,
Margarita [US/US]; Westfield, NJ (US).

(74) Agent: SEIDE, Rochelle, K.; Baker Botts L.L.P., 30
Rockefeller Plaza, New York, NY 10112-4498 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.

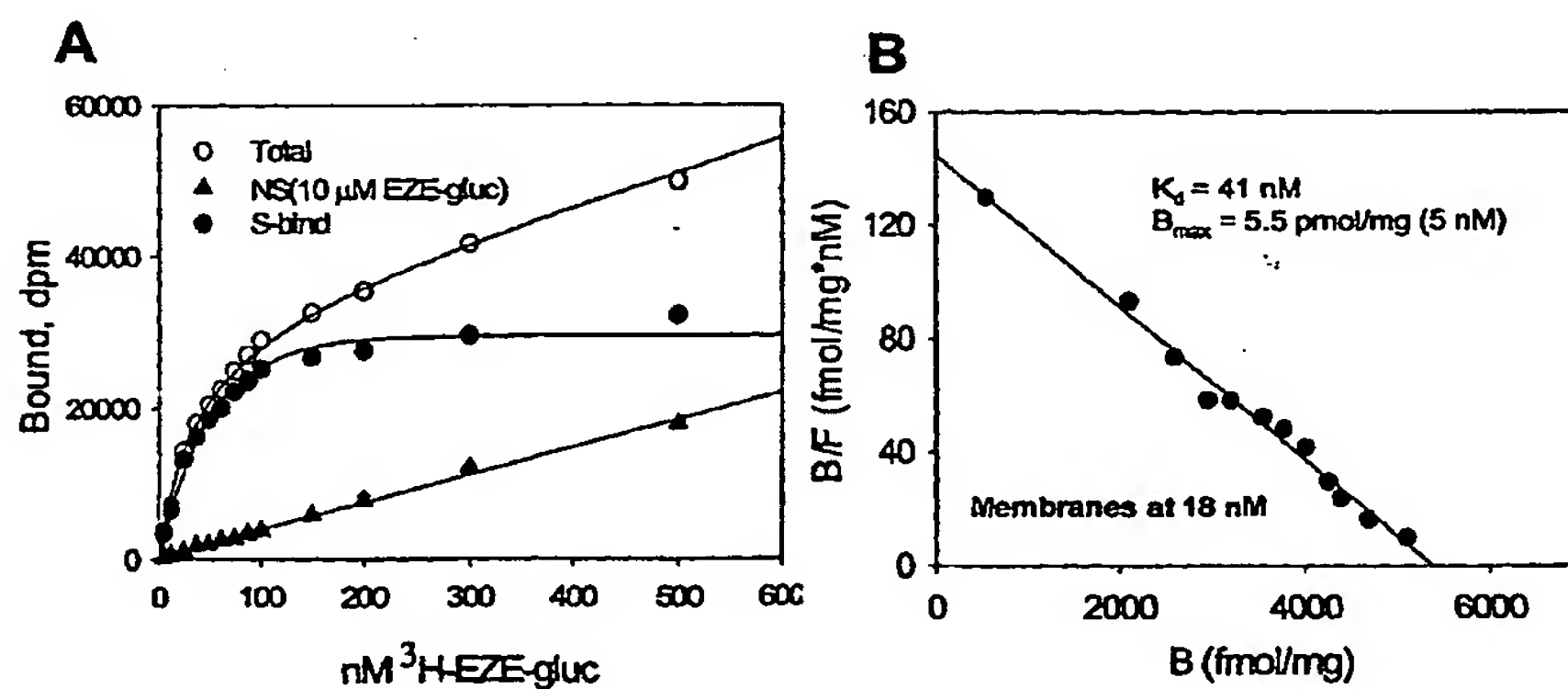
(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO,
SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,
GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF



Equilibrium binding of EZE-glucuronide to rhesus BBMVs

(57) Abstract: The present invention provides human, rat and mouse NPC1L1 polypeptides and polynucleotides encoding the polypeptides. Methods for detecting ligands which bind to NPC1L1 and block intestinal cholesterol absorption are provided. Also included is a method of identifying ligands which bind to NPC1L1 using membranes derived from brush border membrane preparations. Compounds that bind to NPC1L1 can be used for inhibiting intestinal cholesterol absorption in a subject.

WO 2005/069900 A2

NPC1L1 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF

SPECIFICATION

5 The invention claimed herein was made on behalf of Merck & Co., Inc. and Schering-Plough Corporation, parties to a joint research agreement that was in effect on or before the date the claimed invention was made.

 This application claims priority to Serial No. 60/537,341, filed January 16, 2004.

FIELD OF THE INVENTION

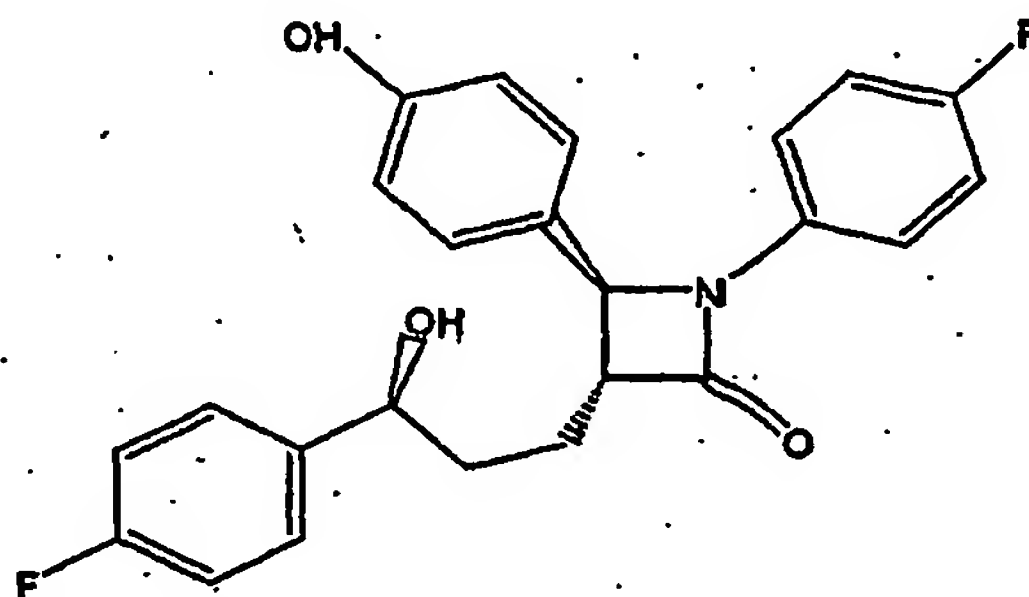
10 The present invention includes NPC1L1 polypeptides and polynucleotides which encode the polypeptides, methods of use and methods of identifying modulators and ligands thereof.

BACKGROUND OF THE INVENTION

15 A factor leading to development of vascular disease, a leading cause of death in industrialized nations, is elevated serum cholesterol. It is estimated that 19% of Americans between the ages of 20 and 74 years of age have high serum cholesterol. The most prevalent form of vascular disease is arteriosclerosis, a condition associated with the thickening and hardening of the arterial wall. Arteriosclerosis of the large
20 vessels is referred to as atherosclerosis. Atherosclerosis is the predominant underlying factor in vascular disorders such as coronary artery disease, aortic aneurysm, arterial disease of the lower extremities and cerebrovascular disease.

 Cholesteryl esters are a major component of atherosclerotic lesions and the major storage form of cholesterol in arterial wall cells. Formation of cholesteryl
25 esters is also a step in the intestinal absorption of dietary cholesterol. Thus, inhibition of cholesteryl ester formation and reduction of serum cholesterol can inhibit the progression of atherosclerotic lesion formation, decrease the accumulation of cholesteryl esters in the arterial wall, and block the intestinal absorption of dietary cholesterol.

The regulation of whole-body cholesterol homeostasis in mammals and animals involves the regulation of intestinal cholesterol absorption, cellular cholesterol trafficking, dietary cholesterol and modulation of cholesterol biosynthesis, bile acid biosynthesis, steroid biosynthesis and the catabolism of the cholesterol-containing plasma lipoproteins. Regulation of intestinal cholesterol absorption has proven to be an effective means by which to regulate serum cholesterol levels. For example, a cholesterol absorption inhibitor, ezetimibe (



), has been shown to be effective in this regard. A pharmaceutical composition containing ezetimibe is commercially available from Merck/Schering-Plough Pharmaceuticals, Inc. under the trade name Zetia®. Identification of a gene target through which ezetimibe acts is important to understanding the process of cholesterol absorption and to the development of other, novel absorption inhibitors. The present invention addresses this need by providing a rat and a mouse homologue of human NPC1L1 (also known as NPC3; Genbank Accession No. AF192522; Davies, *et al.*, (2000) Genomics 65(2): 137-45 and Ioannou, (2000) Mol. Genet. Metab. 71(1-2): 175-81), an ezetimibe target.

NPC1L1 is an N-glycosylated protein comprising a YQRL (SEQ ID NO: 38) motif (i.e., a *trans*-golgi network to plasma membrane transport signal; see Bos, *et al.*, (1993) EMBO J. 12: 2219-2228; Humphrey, *et al.*, (1993) J. Cell. Biol. 120: 1123-1135; Ponnambalam, *et al.*, (1994) J. Cell. Biol. 125: 253-268 and Rothman, *et al.*, (1996) Science 272: 227-234) which exhibits limited tissue distribution and gastrointestinal abundance. Also, the human *NPC1L1* promoter includes a Sterol Regulated Element Binding Protein 1 (SREBP1) binding consensus sequence (Athaniyar, *et al.*, (1998) Proc. Natl. Acad. Sci. USA 95: 4935-4940; Ericsson, *et al.*, (1996) Proc. Natl. Acad. Sci. USA 93: 945-950; Metherall, *et al.*, (1989) J. Biol. Chem. 264: 15634-15641; Smith, *et al.*, (1990) J. Biol. Chem. 265:

2306-2310; Bennett, *et al.*, (1999) J. Biol. Chem. 274: 13025-13032 and Brown, *et al.*, (1997) Cell 89: 331-340). NPC1L1 has 42% amino acid sequence homology to human NPC1 (Genbank Accession No. AF002020), a receptor responsible for Niemann-Pick C1 disease (Carstea, *et al.*, (1997) Science 277: 228-231). Niemann-Pick C1 disease is a rare genetic disorder in humans which results in accumulation of low density lipoprotein (LDL)-derived unesterified cholesterol in lysosomes (Pentchev, *et al.*, (1994) Biochim. Biophys. Acta. 1225: 235-243 and Vanier, *et al.*, (1991) Biochim. Biophys. Acta. 1096: 328-337). In addition, cholesterol accumulates in the *trans*-golgi network of npc1⁻ cells, and relocation of cholesterol, to and from the plasma membrane, is delayed. NPC1 and NPC1L1 each possess 13 transmembrane spanning segments as well as a sterol-sensing domain (SSD). Several other proteins, including HMG-CoA Reductase (HMG-R), Patched (PTC) and Sterol Regulatory Element Binding Protein Cleavage-Activation Protein (SCAP), include an SSD which is involved in sensing cholesterol levels possibly by a mechanism which involves direct cholesterol binding (Gil, *et al.*, (1985) Cell 41: 249-258; Kumagai, *et al.*, (1995) J. Biol. Chem. 270: 19107-19113; Hua, *et al.*, (1996) Cell 87: 415-426; and Radhakrishnan, A., *et al.*, "Direct binding of cholesterol to the purified membrane region of SCAP: Mechanism for a sterol-sensing domain," *Mol. Cell* 15, 259-268 (2004)).

20

SUMMARY OF THE INVENTION

The present invention is based on the discovery that NPC1L1 is the target through which ezetimibe acts, and consequently plays a critical role in the regulation of sterol and 5 α -stanol intestinal transport and absorption, *e.g.* cholesterol absorption. Accordingly, this invention provides for the use of NPC1L1 in an assay for identifying ligands that block NPC1L1-mediated sterol and 5 α -stanol intestinal transport. The present invention provides methods for identifying ligands of NPC1L1 which involve contacting NPC1L1 with a detectably labeled substituted 2-azetidinone, preferably substituted 2-azetidinone-glucuronide, and a candidate compound, and determining whether the candidate compound binds to NPC1L1. The modulation of the binding of the substituted 2-azetidinone to NPC1L1 by the binding of the candidate compound to NPC1L1 indicates that the candidate compound is a ligand that binds to NPC1L1 and is an inhibitor of sterol and 5 α -stanol absorption.

The present invention also provides a method for identifying a ligand of NPC1L1 comprising contacting NPC1L1 with a detectably labeled substituted 2-azetidinone, preferably substituted 2-azetidinone-glucuronide, and measuring the binding of detectably labeled substituted 2-azetidinone to NPC1L1 in the presence and absence of a candidate compound, wherein decreased binding of the detectably labeled substituted 2-azetidinone to the NPC1L1 in the presence of the candidate compound indicates that said candidate compound is a ligand of NPC1L1 and is an inhibitor of sterol and 5 α -stanol absorption.

The present invention also provides for a method for identifying a compound that inhibits intestinal sterol or 5 α -stanol absorption mediated by NPC1L1 involving contacting NPC1L1 with a detectably labeled ligand and the candidate compound and determining whether the candidate compound binds to NPC1L1, wherein binding of said candidate compound to NPC1L1 modulates binding of said ligand to NPC1L1, wherein said modulation indicates that the candidate compound is an intestinal sterol or 5 α -stanol absorption inhibitor.

The present invention provides methods for identifying an ligand of NPC1L1 comprising (a) contacting a host cell (*e.g.*, human embryonic kidney (HEK) 293 cells, chinese hamster ovary (CHO) cell, a J774 cell, a macrophage cell or a Caco2 cell) expressing a polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or SEQ ID NO: 12 or a functional fragment thereof on a cell surface, in the presence of a known amount of a detectably labeled (*e.g.*, with ^3H , ^{14}C , ^{125}I , ^{35}S or fluorescence labeling) substituted azetidinone (*e.g.*, ezetimibe), with a sample to be tested for the presence of an NPC1L1 ligand; and (b) measuring the amount of detectably labeled substituted azetidinone (*e.g.*, ezetimibe) specifically bound to the polypeptide; wherein an NPC1L1 ligand in the sample is identified by measuring substantially reduced binding of the detectably labeled substituted azetidinone (*e.g.*, ezetimibe) to the polypeptide, compared to what would be measured in the absence of such a ligand.

Another method for identifying an ligand of NPC1L1 is also provided. The method comprises (a) placing, in an aqueous suspension, a plurality of support particles, impregnated with a fluorescer (*e.g.*, yttrium silicate, yttrium oxide, diphenyloxazole and polyvinyltoluene), to which a host cell (*e.g.*, chinese hamster ovary (CHO) cell, a J774 cell, a macrophage cell or a Caco2 cell) expressing a

polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or SEQ ID NO: 12 or a functional fragment thereof on a cell surface are attached; (b) adding, to the suspension, a radiolabeled (*e.g.*, with ^3H , ^{14}C or ^{125}I) substituted azetidinone (*e.g.*, ezetimibe) and a sample to be tested for the presence of a ligand, 5 wherein the radiolabel emits radiation energy capable of activating the fluorescer upon the binding of the substituted azetidinone (*e.g.*, ezetimibe) to the polypeptide to produce light energy, whereas radiolabeled substituted azetidinone (*e.g.*, ezetimibe) that does not bind to the polypeptide is, generally, too far removed from the support particles to enable the radioactive energy to activate the fluorescer; and (c) measuring 10 the light energy emitted by the fluorescer in the suspension; wherein an NPC1L1 ligand in the sample is identified by measuring substantially reduced light energy emission, compared to what would be measured in the absence of such a ligand.

Also provided is a method for identifying a ligand of NPC1L1 comprising (a) contacting a host cell (*e.g.*, Chinese hamster ovary (CHO) cell, a J774 15 cell, a macrophage cell or a Caco2 cell) expressing a polypeptide comprising an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or SEQ ID NO: 12 or a functional fragment thereof on a cell surface with detectably labeled (*e.g.*, with ^3H , ^{14}C or ^{125}I) sterol (*e.g.*, cholesterol) or 5 α -stanol and with a sample to be tested for the presence of an ligand; and (b) measuring the amount of detectably labeled sterol (*e.g.*, 20 cholesterol) or 5 α -stanol in the cell; wherein an NPC1L1 antagonist in the sample is identified by measuring substantially reduced detectably labeled sterol (*e.g.*, cholesterol) or 5 α -stanol within the host cell, compared to what would be measured in the absence of such an antagonist and wherein an NPC1L1 agonist in the sample is identified by measuring substantially increased detectably labeled sterol (*e.g.*, 25 cholesterol) or 5 α -stanol within the host cell, compared to what would be measured in the absence of such an agonist.

The present invention includes methods for inhibiting NPC1L1-mediated intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol uptake, in a subject, by administering a substance identified by the screening methods described herein to the 30 subject. Such substances include compounds such as small molecule antagonists of NPC1L1 other than ezetimibe. Also contemplated are methods for antagonizing NPC1L1-mediated sterol (*e.g.*, cholesterol) or 5 α -stanol absorption by administering anti-NPC1L1 antibodies. NPC1L1-mediated absorption of sterol (*e.g.*, cholesterol) or

5α-stanol can also be antagonized by any method which reduces expression of NPC1L1 in an organism. For example, NPC1L1 expression can be reduced by introduction of anti-sense *NPC1L1* mRNA into a cell of an organism or by genetic mutation of the *NPC1L1* gene in an organism (*e.g.*, by complete knockout, disruption, truncation or by introduction of one or more point mutations).

Also included in the present invention is a mutant transgenic mammal (*e.g.*, mouse, rat, dog, rabbit, pig, guinea pig, cat, horse), preferably a mouse comprising a homozygous or heterozygous mutation (*e.g.*, disruption, truncation, one or more point mutations, knock out) of endogenous, chromosomal *NPC1L1* wherein, preferably, the mouse does not produce any functional NPC1L1 protein. Preferably, the mutant mouse, lacking functional NPC1L1, exhibits a reduced level of intestinal sterol (*e.g.*, cholesterol) or 5α-stanol absorption and/or a reduced level of serum sterol (*e.g.*, cholesterol) or 5α-stanol and/or a reduced level of liver sterol (*e.g.*, cholesterol) or 5α-stanol as compared to that of a non-mutant mouse comprising functional NPC1L1. Preferably, in the mutant mouse chromosome, the region of *NPC1L1* (SEQ ID NO: 45) deleted is from nucleotide 790 to nucleotide 998. In one embodiment, *NPC1L1* (SEQ ID NO: 11) is deleted from nucleotide 767 to nucleotide 975. Any offspring or progeny of a parent *NPC1L1* mutant mouse (*i.e.*, *npc1l1*) of the invention which has inherited an *npc1l1* mutant allele is also part of the present invention.

The scope of the present invention also includes a method for screening a sample for an intestinal sterol (*e.g.*, cholesterol) or 5α-stanol absorption antagonist comprising (a) feeding a sterol (*e.g.*, cholesterol) or 5α-stanol-containing substance (*e.g.*, comprising radiolabeled cholesterol, such as ¹⁴C-cholesterol or ³H-cholesterol) to a first and second mouse comprising a functional *NPC1L1* gene and to a third, mutant mouse lacking a functional *NPC1L1*; (b) administering the sample to the first mouse comprising a functional *NPC1L1* but not to the second mouse; (c) measuring the amount of sterol (*e.g.*, cholesterol) or 5α-stanol absorption in the intestine of said first, second and third mouse (*e.g.*, by measuring serum cholesterol); and (d) comparing the levels of intestinal sterol (*e.g.*, cholesterol) or 5α-stanol absorption in each mouse; wherein the sample is determined to contain the intestinal sterol (*e.g.*, cholesterol) or 5α-stanol absorption antagonist when the level of intestinal sterol (*e.g.*, cholesterol) or 5α-stanol absorption in the first mouse and third mouse are

less than the amount of intestinal sterol (e.g., cholesterol) or 5 α -stanol absorption in the second mouse.

The present invention also encompasses a kit comprising (a) a substituted azetidinone (e.g., ezetimibe) in a pharmaceutical dosage form (e.g., a pill or tablet comprising 10 mg substituted azetidinone (e.g., ezetimibe)); and (b) information, for example in the form of an insert, indicating that NPC1L1 is a target of ezetimibe. The kit may also include simvastatin in a pharmaceutical dosage form (e.g., a pill or tablet comprising 5 mg, 10 mg, 20 mg, 40 mg or 80 mg simvastatin). The simvastatin in pharmaceutical dosage form and the ezetimibe in pharmaceutical dosage form can be associated in a single pill or tablet or in separate pills or tablets.

The present invention also provides any isolated mammalian cell (e.g., isolated mouse cell, isolated rat cell or isolated human cell) which lacks a gene which encodes or can produce a functional NPC1L1 polypeptide. The isolated cell can be isolated from a mutant mouse comprising a homozygous mutation of endogenous, chromosomal *NPC1L1* wherein the mouse does not produce any functional NPC1L1 protein. Further, the mutation can be in a gene which when un-mutated encodes an amino acid sequence of SEQ ID NO: 12 (e.g., comprising a nucleotide sequence of SEQ ID NO: 11). The cell can be isolated or derived from duodenum, gall bladder, liver, small intestine or stomach tissue. The cell can be an enterocyte.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1A shows an equilibrium saturation binding plot exhibiting the binding of ³H-EZE-glucuronide to rhesus brush border membrane (BBM) vesicles. Observed total binding (Total) is shown as open circles; nonspecific binding (NS) as triangles, and specific binding (S-bind) as solid circles.

Figure 1B shows a scatchard analysis of ³H-EZE-glucuronide binding to rhesus brush border membrane vesicles.

Figure 2A shows an equilibrium saturation binding plot exhibiting the binding of ³H-EZE-glucuronide (1) to rat brush border membrane vesicles. Observed total binding (open circles) and nonspecific binding (triangles), determined in the presence of 100 μ M unlabeled ezetimibe glucuronide, are included; specific binding (solid circles) was assessed from the difference between total and nonspecific binding.

Binding was measured at 2.5 mg protein/ml in a volume of 100 μ l after 1 hour incubation. Data were fit by nonlinear regression as described in Methods.

Figure 2B shows scatchard analysis of ^3H -EZE-glucuronide binding to rat brush border membrane vesicles. The binding data identify a single high-affinity site with $K_D = 542$ nM and $B_{\text{max}} = 20.7$ pmol/mg protein.

Figure 3A shows association kinetic analysis of ^3H -EZE-glucuronide in rat brush border membrane vesicles. Conditions were 25 nM of 1 and 3 mg/ml protein at 25°C. The second-order rate constant k_{on} ($0.55 \times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$) was calculated from k_{obs} (0.004 s^{-1}) as described in Methods.

Figure 3B shows dissociation kinetic analysis of ^3H -EZE-glucuronide 1 in rat brush border membrane vesicles. After the complex was formed by incubating 25 nM of 1 and 3 mg/ml protein for 1 hour, dissociation was initiated by competition with 100 μ M unlabeled ezetimibe glucuronide. The curve is theoretical for $k_{\text{off}} = 0.0024 \text{ s}^{-1}$.

Figure 4A shows association kinetic analysis of ^3H -EZE-glucuronide in rhesus brush border membrane vesicles.

Figure 4B shows dissociation kinetic analysis of ^3H -EZE-glucuronide in rhesus brush border membrane vesicles.

Figure 5 shows the results of a binding assay where ^3H -EZE-glucuronide is dissociated by EZE-glucuronide and compound 2 from rhesus (A) and rat (B) brush border membrane vesicles.

Figure 6 shows the results of a binding assay where ^{35}S -2 is dissociated by EZE-glucuronide and 2 from mouse brush border membrane vesicles.

Figure 7 shows the distribution of ^3H -EZE-glucuronide binding to rhesus (A) and rat (B) brush border membranes prepared from various portions of rhesus (A) and rat (B) intestinal tissue.

Figure 8 shows the results of a binding assay where ^{35}S -2 is dissociated by EZE-glucuronide and various analogs from CHO cells transfected with rat NPC1L1.

Figure 9 shows the results of a binding assay where ^{35}S -2 is dissociated by EZE-glucuronide and various analogs from CHO cells transfected with human NPC1L1.

Figure 10 shows the binding of ^{35}S -**2** to brush border membrane vesicles prepared from wild type (A) and *NPC1L1* knockout (-/-) mice.

Figure 11 shows the results of a binding assay where ^{35}S -**2** is dissociated by compound **2** from mouse wild type and *NPC1L1* knockout (-/-) brush
5 border membrane vesicles.

Figure 12A shows equilibrium determination of K_D for ezetimibe glucuronide by competition of unlabeled compound against **1** in rat enterocyte brush border membranes. Membranes (1.5 mg/ml protein) were incubated with **1** (50 nM) and the indicated concentrations of ezetimibe glucuronide for 1 hour to ensure
10 equilibrium. K_D at equilibrium is 600 nM. Figure 12B shows the corresponding measurement for rhesus monkey, which were conducted between 0.5 and 1.25 mg/ml protein and 22-50 nM **1**, with incubation time of more than 3 hours. K_D at equilibrium is 38.6 nM.

Figure 13 shows the expression of *NPC1L1* in HEK-293 cells using
15 Western blot analysis (Panel 1) and immunofluorescence (Panel 2).

Figure 14A shows binding of ^3H -ezetimibe glucuronide to enterocyte brush border membranes from wild type mice and *NPC1L1* deficient mice in the presence of detergent. Figure 14B shows competition studies of unlabeled ezetimibe glucuronide against labeled ezetimibe glucuronide.

Figure 15 shows the effect of detergents, taurocholate and digitonin, on
20 [^3H]ezetimibe glucuronide binding.

DETAILED DESCRIPTION OF THE INVENTION

The present invention includes *NPC1L1* polypeptides from rat, human and mouse, along with polynucleotides encoding the respective polypeptides.

25 Preferably, the rat *NPC1L1* polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 2, the human *NPC1L1* comprises the amino acid sequence set forth in SEQ ID NO: 4 and the mouse *NPC1L1* polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 12. The rat *NPC1L1* polynucleotide of SEQ ID NO: 1 or 10 encodes the rat *NPC1L1* polypeptide. The human *NPC1L1*
30 polynucleotide of SEQ ID NO: 3 encodes the human *NPC1L1* polypeptide. The mouse *NPC1L1* polynucleotide of SEQ ID NO: 11 or 13 encodes the mouse *NPC1L1* polypeptide.

The present invention includes any isolated polynucleotide or isolated polypeptide comprising a nucleotide or amino acid sequence referred to, below, in Table 1.

Table 1. Polynucleotides and Polypeptides of the Invention.

Polynucleotide or Polypeptide	Sequence Identifier
Rat <i>NPC1L1</i> polynucleotide	SEQ ID NO: 1
Rat <i>NPC1L1</i> polypeptide	SEQ ID NO: 2
Human <i>NPC1L1</i> polynucleotide	SEQ ID NO: 3
Human <i>NPC1L1</i> polypeptide	SEQ ID NO: 4
Rat <i>NPC1L1</i> expressed sequence tag 603662080F1 (partial sequence)	SEQ ID NO: 5
Rat <i>NPC1L1</i> expressed sequence tag 603665037F1 (partial sequence)	SEQ ID NO: 6
Rat <i>NPC1L1</i> expressed sequence tag 604034587F1 (partial sequence)	SEQ ID NO: 7
EST 603662080F1 with downstream sequences added	SEQ ID NO: 8
EST 603662080F1 with upstream and downstream sequences added	SEQ ID NO: 9
Back-translated polynucleotide sequence of rat NPC1L1	SEQ ID NO: 10
Mouse NPC1L1 polynucleotide	SEQ ID NO: 11
Mouse NPC1L1 polypeptide	SEQ ID NO: 12
Back-translated polynucleotide sequence of mouse NPC1L1	SEQ ID NO: 13
Back-translated polynucleotide sequence of human NPC1L1	SEQ ID NO: 51

5

A human NPC1L1 is also disclosed under Genbank Accession Number AF192522. As discussed below, the nucleotide sequence of the rat *NPC1L1* set forth in SEQ ID NO: 1 was obtained from an expressed sequence tag (EST) from a rat jejunum enterocyte cDNA library. SEQ ID NOs: 5-7 include partial nucleotide

sequences of three independent cDNA clones. The downstream sequence of the SEQ ID NO: 5 EST (603662080F1) were determined; the sequencing data from these experiments are set forth in SEQ ID NO: 8. The upstream sequences were also determined; these data are set forth in SEQ ID NO: 9.

5 SEQ ID NOs: 43 and 44 are the nucleotide and amino acid sequence, respectively, of human NPC1L1 which is disclosed under Genbank Accession No.: AF192522 (see Davies, *et al.*, (2000) Genomics 65(2): 137-45).

 SEQ ID NO: 45 is the nucleotide sequence of a mouse *NPC1L1* which is disclosed under Genbank Accession No. AK078947.

10 NPC1L1 mediates intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption. Inhibition of NPC1L1 in a patient is a useful method for reducing intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption and serum sterol (*e.g.*, cholesterol) or 5 α -stanol in the patient. Reducing the level of intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption and serum sterol (*e.g.*, cholesterol) or 5 α -stanol
15 in a patient is a useful way in which to treat or prevent the occurrence of atherosclerosis, particularly diet-induced atherosclerosis.

 As used herein, the term "sterol" includes, but is not limited to, cholesterol and phytosterols (including, but not limited to, sitosterol, campesterol, stigmasterol and avenosterol).

20 As used herein, the term "5 α -stanol" includes, but is not limited to, cholestanol, 5 α -campestanol and 5 α -sitostanol.

 Without being limited by the present hypothesis, the examples present a better understanding of the putative molecular interaction between NPC1L1 and cholesterol. In this regard, one of the more interesting features of NPC1L1 is that it
25 contains the sterol-sensing domain (SSD) originally observed in SCAP (SREBP cleavage-activating protein). SCAP controls activation of sterol regulatory element binding proteins (SREBP), a transcription factor which controls more than 35 genes related to lipid and cholesterol homeostasis (Brown, M.S. & Goldstein, J.L. A proteolytic pathway that controls the cholesterol content of membranes, cells, and
30 blood. *Proc. Natl. Acad. Sci. U.S.A.* 96, 11041-11048 (1999)). The SSD, consisting of ~180 amino acids in a packet of 5 putative membrane-spanning helices, also serves a regulatory function in two key enzymes on the cholesterol biosynthesis pathway and is present in the receptor Patched. Recently, high affinity binding of cholesterol to the

SSD on SCAP has been demonstrated (Radhakrishnan, A., Sun, L., Kwon, H.J., Brown, M.S. & Goldstein, J.L., "Direct binding of cholesterol to the purified membrane region of SCAP: Mechanism for a sterol-sensing domain," *Mol. Cell* 15, 259-268 (2004)), suggesting that cholesterol may similarly bind to the SSD of
5 NPC1L1, and raising the possibility that ezetimibe may compete with cholesterol for binding at this site.

Molecular Biology

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques
10 within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook, *et al.*, 1989"); DNA Cloning: A Practical Approach, Volumes I and II (D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait, ed.
15 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins, eds. (1985)); Transcription And Translation (B.D. Hames & S.J. Higgins, eds. (1984)); Animal Cell Culture (R.I. Freshney, ed. (1986)); Immobilized Cells And Enzymes (IRL Press, (1986)); B. Perbal, A Practical Guide To Molecular Cloning (1984); F.M. Ausubel, *et al.* (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

20 The back-translated sequences of SEQ ID NO: 10 and of SEQ ID NO: 13 uses the single-letter code shown in Table 1 of Annex C, Appendix 2 of the PCT Administrative Instruction in the Manual of Patent Examination Procedure.

A "polynucleotide", "nucleic acid" or "nucleic acid molecule" may refer to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine,
25 uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules"), or any phosphoester analogs thereof, such as phosphorothioates and thioesters, in single stranded form, double-stranded form or otherwise.

A "polynucleotide sequence", "nucleic acid sequence" or "nucleotide
30 sequence" is a series of nucleotide bases (also called "nucleotides") in a nucleic acid, such as DNA or RNA, and means any chain of two or more nucleotides.

A "coding sequence" or a sequence "encoding" an expression product, such as a RNA, polypeptide, protein, or enzyme, is a nucleotide sequence that, when expressed, results in production of the product.

The term "gene" means a DNA sequence that codes for or corresponds to a particular sequence of ribonucleotides or amino acids which comprise all or part of one or more RNA molecules, proteins or enzymes, and may or may not include regulatory DNA sequences, such as promoter sequences, which determine, for example, the conditions under which the gene is expressed. Genes may be transcribed from DNA to RNA which may or may not be translated into an amino acid sequence.

The present invention includes nucleic acid fragments of any of SEQ ID NOs: 1, 5-11 or 13. A nucleic acid "fragment" includes at least about 30 (*e.g.*, 31, 32, 33, 34), preferably at least about 35 (*e.g.*, 25, 26, 27, 28, 29, 30, 31, 32, 33 or 34), more preferably at least about 45 (*e.g.*, 35, 36, 37, 38, 39, 40, 41, 42, 43 or 44), and most preferably at least about 126 or more contiguous nucleotides (*e.g.*, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 150, 160, 170, 180, 190, 200, 300, 400, 500, 1000 or 1200) from any of SEQ ID NOs: 1, 5-11 or 13.

The present invention also includes nucleic acid fragments consisting of at least about 7 (*e.g.*, 9, 12, 17, 19), preferably at least about 20 (*e.g.*, 30, 40, 50, 60), more preferably about 70 (*e.g.*, 80, 90, 95), yet more preferably at least about 100 (*e.g.*, 105, 110, 114) and even more preferably at least about 115 (*e.g.*, 117, 119, 120, 122, 124, 125, 126) contiguous nucleotides from any of SEQ ID NOs: 1, 5-11 or 13.

As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of no more than about 100 nucleotides (*e.g.*, 30, 40, 50, 60, 70, 80, or 90), that may be hybridizable to a genomic DNA molecule, a cDNA molecule, or an mRNA molecule encoding a gene, mRNA, cDNA, or other nucleic acid of interest. Oligonucleotides can be labeled, *e.g.*, by incorporation of ³²P-nucleotides, ³H-nucleotides, ¹⁴C-nucleotides, ³⁵S-nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of the gene, or to detect the presence of nucleic acids. Generally, oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer.

A "protein sequence", "peptide sequence" or "polypeptide sequence" or "amino acid sequence" may refer to a series of two or more amino acids in a protein, peptide or polypeptide.

5 "Protein", "peptide" or "polypeptide" includes a contiguous string of two or more amino acids. Preferred peptides of the invention include those set forth in any of SEQ ID NOs: 2 or 12 as well as variants and fragments thereof. Such fragments preferably comprise at least about 10 (*e.g.*, 11, 12, 13, 14, 15, 16, 17, 18 or 19), more preferably at least about 20 (*e.g.*, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40), and yet more preferably at least about 42 (*e.g.*, 43, 44, 45, 46, 47, 48, 49, 50, 60, 10 70, 80, 90, 100, 110, 120 or 130) or more contiguous amino acid residues from any of SEQ ID NOs: 2 or 12.

The present invention also includes polypeptides, preferably antigenic polypeptides, consisting of at least about 7 (*e.g.*, 9, 10, 13, 15, 17, 19), preferably at least about 20 (*e.g.*, 22, 24, 26, 28), yet more preferably at least about 30 (*e.g.*, 32, 34, 15 36, 38) and even more preferably at least about 40 (*e.g.*, 41, 42) contiguous amino acids from any of SEQ ID NOs: 2 or 12.

The polypeptides of the invention can be produced by proteolytic cleavage of an intact peptide, by chemical synthesis or by the application of recombinant DNA technology and are not limited to polypeptides delineated by 20 proteolytic cleavage sites. The polypeptides, either alone or cross-linked or conjugated to a carrier molecule to render them more immunogenic, are useful as antigens to elicit the production of antibodies and fragments thereof. The antibodies can be used, *e.g.*, in immunoassays for immunoaffinity purification or for inhibition of NPC1L1, etc.

25 The terms "isolated polynucleotide" or "isolated polypeptide" include a polynucleotide (*e.g.*, RNA or DNA molecule, or a mixed polymer) or a polypeptide, respectively, which are partially or fully separated from other components that are normally found in cells or in recombinant DNA expression systems. These components include, but are not limited to, cell membranes, cell walls, ribosomes, 30 polymerases, serum components and extraneous genomic sequences.

An isolated polynucleotide or polypeptide will, preferably, be an essentially homogeneous composition of molecules but may contain some heterogeneity.

“Amplification” of DNA as used herein may denote the use of polymerase chain reaction (PCR) to increase the concentration of a particular DNA sequence within a mixture of DNA sequences. For a description of PCR see Saiki, *et al.*, Science (1988) 239: 487.

5 The term “host cell” includes any cell of any organism that is selected, modified, transfected, transformed, grown, or used or manipulated in any way, for the production of a substance by the cell, for example, the expression or replication, by the cell, of a gene, a DNA or RNA sequence or a protein. Preferred host cells include HEK-293 cells, chinese hamster ovary (CHO) cells, murine macrophage J774 cells or
10 any other macrophage cell line and human intestinal epithelial Caco2 cells.

 The nucleotide sequence of a nucleic acid may be determined by any method known in the art (*e.g.*, chemical sequencing or enzymatic sequencing). “Chemical sequencing” of DNA includes methods such as that of Maxam and Gilbert (1977) (Proc. Natl. Acad. Sci. USA 74: 560), in which DNA is randomly cleaved
15 using individual base-specific reactions. “Enzymatic sequencing” of DNA includes methods such as that of Sanger (Sanger, *et al.*, (1977) Proc. Natl. Acad. Sci. USA 74: 5463).

 The nucleic acids herein may be flanked by natural regulatory (expression control) sequences, or may be associated with heterologous sequences,
20 including promoters, internal ribosome entry sites (IRES) and other ribosome binding site sequences, enhancers, response elements, suppressors, signal sequences, polyadenylation sequences, introns, 5'- and 3'- non-coding regions, and the like.

 In general, a “promoter” or “promoter sequence” is a DNA regulatory region capable of binding an RNA polymerase in a cell (*e.g.*, directly or through other
25 promoter-bound proteins or substances) and initiating transcription of a coding sequence. A promoter sequence is, in general, bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at any level. Within the promoter sequence may be found a transcription initiation site
30 (conveniently defined, for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. The promoter may be operably associated with other expression control sequences, including enhancer and repressor sequences or with a nucleic acid of the

invention. Promoters which may be used to control gene expression include, but are not limited to, cytomegalovirus (CMV) promoter (U.S. Patent Nos. 5,385,839 and 5,168,062), the SV40 early promoter region (Benoist, *et al.*, (1981) *Nature* 290: 304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus
5 (Yamamoto, *et al.*, (1980) *Cell* 22: 787-797), the herpes thymidine kinase promoter (Wagner, *et al.*, (1981) *Proc. Natl. Acad. Sci. USA* 78: 1441-1445), the regulatory sequences of the metallothionein gene (Brinster, *et al.*, (1982) *Nature* 296: 39-42); prokaryotic expression vectors such as the β -lactamase promoter (Villa-Komaroff, *et al.*, (1978) *Proc. Natl. Acad. Sci. USA* 75: 3727-3731), or the *tac* promoter
10 (DeBoer, *et al.*, (1983) *Proc. Natl. Acad. Sci. USA* 80: 21-25); see also "Useful proteins from recombinant bacteria" in *Scientific American* (1980) 242: 74-94; and promoter elements from yeast or other fungi such as the *Gal 4* promoter, the *ADC* (alcohol dehydrogenase) promoter, *PGK* (phosphoglycerol kinase) promoter or the alkaline phosphatase promoter.

15 A coding sequence is "under the control of", "functionally associated with" or "operably associated with" transcriptional and translational control sequences in a cell when the sequences direct RNA polymerase mediated transcription of the coding sequence into RNA, preferably mRNA, which then may be RNA spliced (if it contains introns) and, optionally, translated into a protein encoded
20 by the coding sequence.

The terms "express" and "expression" mean allowing or causing the information in a gene, RNA or DNA sequence to become manifest; for example, producing a protein by activating the cellular functions involved in transcription and translation of a corresponding gene. A DNA sequence is expressed in or by a cell to
25 form an "expression product" such as an RNA (*e.g.*, mRNA) or a protein. The expression product itself may also be said to be "expressed" by the cell.

The term "transformation" means the introduction of a nucleic acid into a cell. The introduced gene or sequence may be called a "clone". A host cell that receives the introduced DNA or RNA has been "transformed" and is a "transformant"
30 or a "clone." The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell, or from cells of a different genus or species.

The term "vector" includes a vehicle (e.g., a plasmid) by which a DNA or RNA sequence can be introduced into a host cell, so as to transform the host and, optionally, promote expression and/or replication of the introduced sequence.

Vectors that can be used in this invention include plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles that may facilitate introduction of the nucleic acids into the genome of the host. Plasmids are the most commonly used form of vector but all other forms of vectors which serve a similar function and which are, or become, known in the art are suitable for use herein. See, e.g., Pouwels, *et al.*, Cloning Vectors: A Laboratory Manual, 1985 and Supplements, Elsevier, N.Y., and Rodriguez *et al.* (eds.), Vectors: A Survey of Molecular Cloning Vectors and Their Uses, 1988, Butterworth, Boston, MA.

The term "expression system" means a host cell and compatible vector which, under suitable conditions, can express a protein or nucleic acid which is carried by the vector and introduced to the host cell. Common expression systems include *E. coli* host cells and plasmid vectors, insect host cells and Baculovirus vectors, and mammalian host cells and vectors.

Expression of nucleic acids encoding the NPC1L1 polypeptides of this invention can be carried out by conventional methods in either prokaryotic or eukaryotic cells. Although *E. coli* host cells are employed most frequently in prokaryotic systems, many other bacteria, such as various strains of *Pseudomonas* and *Bacillus*, are known in the art and can be used as well. Suitable host cells for expressing nucleic acids encoding the NPC1L1 polypeptides include prokaryotes and higher eukaryotes. Prokaryotes include both gram-negative and gram-positive organisms, e.g., *E. coli* and *B. subtilis*. Higher eukaryotes include established tissue culture cell lines from animal cells, both of non-mammalian origin, e.g., insect cells, and birds, and of mammalian origin, e.g., human, primates, and rodents.

Prokaryotic host-vector systems include a wide variety of vectors for many different species. A representative vector for amplifying DNA is pBR322 or many of its derivatives (e.g., pUC18 or 19). Vectors that can be used to express the NPC1L1 polypeptides include, but are not limited to, those containing the *lac* promoter (pUC-series); *trp* promoter (pBR322-*trp*); *Ipp* promoter (the pIN-series); lambda-pP or pR promoters (pOTS); or hybrid promoters such as *ptac* (pDR540). See Brosius *et al.*, "Expression Vectors Employing Lambda-, *trp*-, *lac*-, and *Ipp*-derived

Promoters", in Rodriguez and Denhardt (eds.) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, 1988, Butterworth, Boston, pp. 205-236. Many polypeptides can be expressed, at high levels, in an *E.coli*/T7 expression system as disclosed in U.S. Patent Nos. 4,952,496; 5,693,489 and 5,869,320 and in Davanloo, P., *et al.*, (1984) Proc. Natl. Acad. Sci. USA 81: 2035-2039; Studier, F.W., *et al.*, (1986) J. Mol. Biol. 189: 113-130; Rosenberg, A. H., *et al.*, (1987) Gene 56: 125-135; and Dunn, J.J., *et al.*, (1988) Gene 68: 259.

Higher eukaryotic tissue culture cells may also be used for the recombinant production of the NPC1L1 polypeptides of the invention. Although any higher eukaryotic tissue culture cell line might be used, including insect baculovirus expression systems, mammalian cells are preferred. Transformation or transfection and propagation of such cells have become a routine procedure. Examples of useful cell lines include HeLa cells, chinese hamster ovary (CHO) cell lines, J774 cells, HEK-293 cells, Caco2 cells, baby rat kidney (BRK) cell lines, insect cell lines, bird cell lines, and monkey (COS) cell lines. Expression vectors for such cell lines usually include an origin of replication, a promoter, a translation initiation site, RNA splice sites (if genomic DNA is used), a polyadenylation site, and a transcription termination site. These vectors also, usually, contain a selection gene or amplification gene. Suitable expression vectors may be plasmids, viruses, or retroviruses carrying promoters derived, *e.g.*, from such sources as adenovirus, SV40, parvoviruses, vaccinia virus, or cytomegalovirus. Examples of expression vectors include pCR@3.1, pCDNA1, pCD (Okayama, *et al.*, (1985) Mol. Cell Biol. 5: 1136), pMC1neo Poly-A (Thomas, *et al.*, (1987) Cell 51: 503), pREP8, pSVSPORT and derivatives thereof, and baculovirus vectors such as pAC373 or pAC610. One embodiment of the invention includes membrane bound NPC1L1. In this embodiment, NPC1L1 can be expressed in the cell membrane of a eukaryotic cell and the membrane bound protein can be isolated from the cell by conventional methods which are known in the art.

The present invention also includes fusions which include the NPC1L1 polypeptides and *NPC1L1* polynucleotides of the present invention and a second polypeptide or polynucleotide moiety, which may be referred to as a "tag". The fusions of the present invention may comprise any of the polynucleotides or polypeptides set forth in Table 1 or any subsequence or fragment thereof (discussed

above). The fused polypeptides of the invention may be conveniently constructed, for example, by insertion of a polynucleotide of the invention or fragment thereof into an expression vector. The fusions of the invention may include tags which facilitate purification or detection. Such tags include glutathione-S-transferase (GST),
5 hexahistidine (His6) tags, maltose binding protein (MBP) tags, haemagglutinin (HA) tags, cellulose binding protein (CBP) tags and myc tags. Detectable tags such as ^{32}P , ^{35}S , ^3H , $^{99\text{m}}\text{Tc}$, ^{123}I , ^{111}In , ^{68}Ga , ^{18}F , ^{125}I , ^{131}I , $^{113\text{m}}\text{In}$, ^{76}Br , ^{67}Ga , $^{99\text{m}}\text{Tc}$, ^{123}I , ^{111}In and ^{68}Ga may also be used to label the polypeptides and polynucleotides of the invention. Methods for constructing and using such fusions are very conventional and well
10 known in the art.

Modifications (*e.g.*, post-translational modifications) that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications, in large part, will be determined by the host cell's post-translational
15 modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as *E. coli*. Accordingly, when glycosylation is desired, a polypeptide can be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out post-translational glycosylations which are similar to those
20 of mammalian cells. For this reason, insect cell expression systems have been developed to express, efficiently, mammalian proteins having native patterns of glycosylation. An insect cell which may be used in this invention is any cell derived from an organism of the class *Insecta*. Preferably, the insect is *Spodoptera frugiperda* (Sf9 or Sf21) or *Trichoplusia ni* (High 5). Examples of insect expression
25 systems that can be used with the present invention, for example to produce NPC1L1 polypeptide, include Bac-To-Bac (Invitrogen Corporation, Carlsbad, CA) or Gateway (Invitrogen Corporation, Carlsbad, CA). If desired, deglycosylation enzymes can be used to remove carbohydrates attached during production in eukaryotic expression systems.

30 Other modifications may also include addition of aliphatic esters or amides to the polypeptide carboxyl terminus. The present invention also includes analogs of the NPC1L1 polypeptides which contain modifications, such as incorporation of unnatural amino acid residues, or phosphorylated amino acid

residues such as phosphotyrosine, phosphoserine or phosphothreonine residues.

Other potential modifications include sulfonation, biotinylation, or the addition of other moieties. For example, the NPC1L1 polypeptides of the invention may be appended with a polymer which increases the half-life of the peptide in the body of a
5 subject. Preferred polymers include polyethylene glycol (PEG) (*e.g.*, PEG with a molecular weight of 2 kDa, 5 kDa, 10 kDa, 12 kDa, 20 kDa, 30 kDa and 40 kDa), dextran and monomethoxypolyethylene glycol (mPEG).

The peptides of the invention may also be cyclized. Specifically, the amino- and carboxy-terminal residues of an NPC1L1 polypeptide or two internal
10 residues of an NPC1L1 polypeptide of the invention can be fused to create a cyclized peptide. Methods for cyclizing peptides are conventional and very well known in the art; for example, see Gurrath, *et al.*, (1992) *Eur. J. Biochem.* 210: 911-921.

The present invention contemplates any superficial or slight modification to the amino acid or nucleotide sequences which correspond to the
15 polypeptides of the invention. In particular, the present invention contemplates sequence conservative variants of the nucleic acids which encode the polypeptides of the invention. "Sequence-conservative variants" of a polynucleotide sequence are those in which a change of one or more nucleotides in a given codon results in no alteration in the amino acid encoded at that position. Function-conservative variants
20 of the polypeptides of the invention are also contemplated by the present invention. "Function-conservative variants" are those in which one or more amino acid residues in a protein or enzyme have been changed without altering the overall conformation and function of the polypeptide, including, but, by no means, limited to, replacement of an amino acid with one having similar properties. Amino acids with similar
25 properties are well known in the art. For example, polar/hydrophilic amino acids which may be interchangeable include asparagine, glutamine, serine, cysteine, threonine, lysine, arginine, histidine, aspartic acid and glutamic acid; nonpolar/hydrophobic amino acids which may be interchangeable include glycine, alanine, valine, leucine, isoleucine, proline, tyrosine, phenylalanine, tryptophan and
30 methionine; acidic amino acids, which may be interchangeable include aspartic acid and glutamic acid and basic amino acids, which may be interchangeable include histidine, lysine and arginine.

The present invention includes polynucleotides encoding rat, human or mouse NPC1L1 and fragments thereof as well as nucleic acids which hybridize to the polynucleotides. Preferably, the nucleic acids hybridize under low stringency conditions, more preferably under moderate stringency conditions and most preferably under high stringency conditions. A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (see Sambrook, *et al.*, *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. Typical low stringency hybridization conditions are 55°C, 5X SSC, 0.1% SDS, 0.25% milk, and no formamide at 42°C; or 30% formamide, 5X SSC, 0.5% SDS at 42°C. Typical, moderate stringency hybridization conditions are similar to the low stringency conditions except the hybridization is carried out in 40% formamide, with 5X or 6X SSC at 42°C. High stringency hybridization conditions are similar to low stringency conditions except the hybridization conditions are carried out in 50% formamide, 5X or 6X SSC and, optionally, at a higher temperature (*e.g.*, higher than 42°C: 57°C, 59°C, 60°C, 62°C, 63°C, 65°C or 68°C). In general, SSC is 0.15M NaCl and 0.015M Na-citrate. Hybridization requires that the two nucleic acids contain complementary sequences, although, depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the higher the stringency under which the nucleic acids may hybridize. For hybrids of greater than 100 nucleotides in length, equations for calculating the melting temperature have been derived (see Sambrook, *et al.*, *supra*, 9.50-9.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook, *et al.*, *supra*).

Also included in the present invention are polynucleotides comprising nucleotide sequences and polypeptides comprising amino acid sequences which are at least about 70% identical, preferably at least about 80% identical, more preferably at least about 90% identical and most preferably at least about 95% identical (*e.g.*, 95%,

96%, 97%, 98%, 99%, 100%) to the reference rat *NPC1L1* nucleotide (*e.g.*, any of SEQ ID NOs: 1 or 5-10) and amino acid sequences (*e.g.*, SEQ ID NO: 2), reference human *NPC1L1* nucleotide (*e.g.*, SEQ ID NO: 3) and amino acid sequences (*e.g.*, SEQ ID NO: 4) or the reference mouse *NPC1L1* nucleotide (*e.g.*, any of SEQ ID
5 NOs: 11 or 13) and amino acid sequences (*e.g.*, SEQ ID NO: 12), when the comparison is performed by a BLAST algorithm wherein the parameters of the algorithm are selected to give the largest match between the respective sequences over the entire length of the respective reference sequences. Polypeptides comprising amino acid sequences which are at least about 70% similar, preferably at least about
10 80% similar, more preferably at least about 90% similar and most preferably at least about 95% similar (*e.g.*, 95%, 96%, 97%, 98%, 99%, 100%) to the reference rat NPC1L1 amino acid sequence of SEQ ID NO: 2, reference human NPC1L1 amino acid sequence of SEQ ID NO: 4 or the reference mouse NPC1L1 amino acid sequence of SEQ ID NO: 12, when the comparison is performed with a BLAST algorithm
15 wherein the parameters of the algorithm are selected to give the largest match between the respective sequences over the entire length of the respective reference sequences, are also included in the present invention.

Sequence identity refers to exact matches between the nucleotides or amino acids of two sequences which are being compared. Sequence similarity refers
20 to both exact matches between the amino acids of two polypeptides which are being compared in addition to matches between nonidentical, biochemically related amino acids. Biochemically related amino acids which share similar properties and may be interchangeable are discussed above.

The following references regarding the BLAST algorithm are herein
25 incorporated by reference: **BLAST ALGORITHMS**: Altschul, S.F., *et al.*, (1990) J. Mol. Biol. 215: 403-410; Gish, W., *et al.*, (1993) Nature Genet. 3: 266-272; Madden, T.L., *et al.*, (1996) Meth. Enzymol. 266: 131-141; Altschul, S.F., *et al.*, (1997) Nucleic Acids Res. 25: 3389-3402; Zhang, J., *et al.*, (1997) Genome Res. 7: 649-656; Wootton, J.C., *et al.*, (1993) Comput. Chem. 17: 149-163; Hancock, J.M., *et al.*,
30 (1994) Comput. Appl. Biosci. 10: 67-70; **ALIGNMENT SCORING SYSTEMS**: Dayhoff, M.O., *et al.*, "A model of evolutionary change in proteins" in Atlas of Protein Sequence and Structure, (1978) vol. 5, suppl. 3. M.O. Dayhoff (ed.), pp. 345-352, Natl. Biomed. Res. Found., Washington, DC; Schwartz, R.M., *et al.*, "Matrices

for detecting distant relationships" in Atlas of Protein Sequence and Structure, (1978) vol. 5, suppl. 3. M.O. Dayhoff (ed.), pp. 353-358, Natl. Biomed. Res. Found., Washington, DC; Altschul, S.F., (1991) J. Mol. Biol. 219: 555-565; States, D.J., *et al.*, (1991) Methods 3: 66-70; Henikoff, S., *et al.*, (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919; Altschul, S.F., *et al.*, (1993) J. Mol. Evol. 36: 290-300; **ALIGNMENT STATISTICS**: Karlin, S., *et al.*, (1990) Proc. Natl. Acad. Sci. USA 87: 2264-2268; Karlin, S., *et al.*, (1993) Proc. Natl. Acad. Sci. USA 90: 5873-5877; Dembo, A., *et al.*, (1994) Ann. Prob. 22: 2022-2039; and Altschul, S.F. "Evaluating the statistical significance of multiple distinct local alignments" in Theoretical and Computational Methods in Genome Research (S. Suhai, ed.), (1997) pp. 1-14, Plenum, New York.

Protein Purification

The proteins, polypeptides and antigenic fragments of this invention can be purified by standard methods, including, but not limited to, salt or alcohol precipitation, affinity chromatography (*e.g.*, used in conjunction with a purification tagged NPC1L1 polypeptide as discussed above), preparative disc-gel electrophoresis, isoelectric focusing, high pressure liquid chromatography (HPLC), reversed-phase HPLC, gel filtration, cation and anion exchange and partition chromatography, and countercurrent distribution. Such purification methods are well known in the art and are disclosed, *e.g.*, in "Guide to Protein Purification", Methods in Enzymology, Vol. 182, M. Deutscher, Ed., 1990, Academic Press, New York, NY.

Purification steps can be followed by performance of assays for receptor binding activity as described below. Particularly where an NPC1L1 polypeptide is being isolated from a cellular or tissue source, it is preferable to include one or more inhibitors of proteolytic enzymes in the assay system, such as phenylmethanesulfonyl fluoride (PMSF), Pefabloc SC, pepstatin, leupeptin, chymostatin and EDTA.

Antibody Molecules

Antigenic (including immunogenic) fragments of the NPC1L1 polypeptides of the invention are within the scope of the present invention (*e.g.*, 42 or more contiguous amino acids from SEQ ID NO: 2, 4 or 12). The antigenic peptides may be useful, *inter alia*, for preparing isolated antibody molecules which recognize NPC1L1. Isolated anti-NPC1L1 antibody molecules are useful NPC1L1 ligands.

An antigen is any molecule that can bind specifically to an antibody. Some antigens cannot, by themselves, elicit antibody production. Those that can induce antibody production are immunogens.

Preferably, isolated anti-NPC1L1 antibodies recognize an antigenic peptide comprising an amino acid sequence selected from SEQ ID NOs: 39-42 (*e.g.*, an antigen derived from rat NPC1L1). More preferably, the antibody is A0715, A0716, A0717, A0718, A0867, A0868, A1801 or A1802.

The term "antibody molecule" includes, but is not limited to, antibodies and fragments (preferably antigen-binding fragments) thereof. The term includes monoclonal antibodies, polyclonal antibodies, bispecific antibodies, Fab antibody fragments, F(ab)₂ antibody fragments, Fv antibody fragments (*e.g.*, V_H or V_L), single chain Fv antibody fragments and dsFv antibody fragments. Furthermore, the antibody molecules of the invention may be fully human antibodies, mouse antibodies, rat antibodies, rabbit antibodies, goat antibodies, chicken antibodies, humanized antibodies or chimeric antibodies.

Although it is not always necessary, when NPC1L1 polypeptides are used as antigens to elicit antibody production in an immunologically competent host, smaller antigenic fragments are, preferably, first rendered more immunogenic by cross-linking or concatenation, or by coupling to an immunogenic carrier molecule (*i.e.*, a macromolecule having the property of independently eliciting an immunological response in a host animal, such as diphtheria toxin or tetanus). Cross-linking or conjugation to a carrier molecule may be required because small polypeptide fragments sometimes act as haptens (molecules which are capable of specifically binding to an antibody but incapable of eliciting antibody production, *i.e.*, they are not immunogenic). Conjugation of such fragments to an immunogenic carrier molecule renders them more immunogenic through what is commonly known as the "carrier effect".

Carrier molecules include, *e.g.*, proteins and natural or synthetic polymeric compounds such as polypeptides, polysaccharides, lipopolysaccharides, etc. Protein carrier molecules are especially preferred, including, but not limited to, keyhole limpet hemocyanin and mammalian serum proteins such as human or bovine gammaglobulin, human, bovine or rabbit serum albumin, or methylated or other derivatives of such proteins. Other protein carriers will be apparent to those skilled in

the art. Preferably, the protein carrier will be foreign to the host animal in which antibodies against the fragments are to be elicited.

Covalent coupling to the carrier molecule can be achieved using methods well known in the art, the exact choice of which will be dictated by the nature of the carrier molecule used. When the immunogenic carrier molecule is a protein, the fragments of the invention can be coupled, *e.g.*, using water-soluble carbodiimides such as dicyclohexylcarbodiimide or glutaraldehyde.

Coupling agents, such as these, can also be used to cross-link the fragments to themselves without the use of a separate carrier molecule. Such cross-linking into aggregates can also increase immunogenicity. Immunogenicity can also be increased by the use of known adjuvants, alone or in combination with coupling or aggregation.

Adjuvants for the vaccination of animals include, but are not limited to, Adjuvant 65 (containing peanut oil, mannide monooleate and aluminum monostearate); Freund's complete or incomplete adjuvant; mineral gels such as aluminum hydroxide, aluminum phosphate and alum; surfactants such as hexadecylamine, octadecylamine, lysolecithin, dimethyldioctadecylammonium bromide, N,N-dioctadecyl-N',N'-bis(2-hydroxymethyl) propanediamine, methoxyhexadecylglycerol and pluronic polyols; polyanions such as pyran, dextran sulfate, poly IC, polyacrylic acid and carbopol; peptides such as muramyl dipeptide, dimethylglycine and tuftsin; and oil emulsions. The polypeptides could also be administered following incorporation into liposomes or other microcarriers.

Information concerning adjuvants and various aspects of immunoassays are disclosed, *e.g.*, in the series by P. Tijssen, Practice and Theory of Enzyme Immunoassays, 3rd Edition, 1987, Elsevier, New York. Other useful references covering methods for preparing polyclonal antisera include Microbiology, 1969, Hoeber Medical Division, Harper and Row; Landsteiner, Specificity of Serological Reactions, 1962, Dover Publications, New York, and Williams, *et al.*, Methods in Immunology and Immunochemistry, Vol. 1, 1967, Academic Press, New York.

The anti-NPC1L1 antibody molecules of the invention preferably recognize human, mouse or rat NPC1L1; however, the present invention includes antibody molecules which recognize NPC1L1 from any species, preferably mammals

(e.g., cat, sheep or horse). The present invention also includes complexes comprising an NPC1L1 polypeptide of the invention and an anti-NPC1L1 antibody molecule. Such complexes can be made by simply contacting the antibody molecule with its cognate polypeptide.

5 Various methods may be used to make the antibody molecules of the invention. Human antibodies can be made, for example, by methods which are similar to those disclosed in U.S. Patent Nos. 5,625,126; 5,877,397; 6,255,458; 6,023,010 and 5,874,299.

10 Hybridoma cells which produce the monoclonal anti-NPC1L1 antibodies may be produced by methods which are commonly known in the art. These methods include, but are not limited to, the hybridoma technique originally developed by Kohler, *et al.*, (1975) (Nature 256: 495-497), as well as the trioma technique (Hering, *et al.*, (1988) Biomed. Biochim. Acta. 47: 211-216 and Hagiwara, *et al.*, (1993) Hum. Antibod. Hybridomas 4: 15), the human B-cell hybridoma
15 technique (Kozbor, *et al.*, (1983) Immunology Today 4: 72 and Cote, *et al.*, (1983) Proc. Natl. Acad. Sci. U.S.A. 80: 2026-2030), and the EBV-hybridoma technique (Cole, *et al.*, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96, 1985). ELISA may be used to determine if hybridoma cells are expressing anti-NPC1L1 antibodies.

20 The anti-NPC1L1 antibody molecules of the present invention may also be produced recombinantly (e.g., in an *E.coli*/T7 expression system as discussed above). In this embodiment, nucleic acids encoding the antibody molecules of the invention (e.g., V_H or V_L) may be inserted into a pet-based plasmid and expressed in the *E.coli*/T7 system. There are several methods by which to produce recombinant
25 antibodies which are known in the art. An example of a method for recombinant production of antibodies is disclosed in U.S. Patent No. 4,816,567. See also Skerra, A., *et al.*, (1988) Science 240: 1038-1041; Better, M., *et al.*, (1988) Science 240: 1041-1043 and Bird, R.E., *et al.*, (1988) Science 242: 423-426.

30 The term "monoclonal antibody," includes an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible, naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Monoclonal antibodies are

advantageous in that they may be synthesized by a hybridoma culture, essentially uncontaminated by other immunoglobulins. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. The monoclonal antibodies to be used in
5 accordance with the present invention may be made by the hybridoma method as described by Kohler, *et al.*, (1975) Nature 256: 495.

The term "polyclonal antibody" includes an antibody which was produced among or in the presence of one or more other, non-identical antibodies. In
10 general, polyclonal antibodies are produced from a B-lymphocyte in the presence of several other B-lymphocytes which produced non-identical antibodies. Typically, polyclonal antibodies are obtained directly from an immunized animal (*e.g.*, a rabbit).

A "bispecific antibody" comprises two different antigen binding regions which bind to distinct antigens. Bispecific antibodies, as well as methods of
15 making and using the antibodies, are conventional and very well known in the art.

Anti-idiotypic antibodies or anti-idiotypes are antibodies directed against the antigen-combining region or variable region (called the idiotype) of another antibody molecule. As disclosed by Jerne (Jerne, N. K., (1974) Ann. Immunol. (Paris) 125c: 373 and Jerne, N. K., *et al.*, (1982) EMBO 1: 234),
20 immunization with an antibody molecule expressing a paratope (antigen-combining site) for a given antigen (*e.g.*, NPC1L1) will produce a group of anti-antibodies, some of which share, with the antigen, a complementary structure to the paratope. Immunization with a subpopulation of the anti-idiotypic antibodies will, in turn, produce a subpopulation of antibodies or immune cell subsets that are reactive to the
25 initial antigen.

The term "fully human antibody" refers to an antibody which comprises human immunoglobulin sequences only. Similarly, "mouse antibody" refers to an antibody which comprises mouse immunoglobulin sequences only.

"Human/mouse chimeric antibody" refers to an antibody which
30 comprises a mouse variable region (V_H and V_L) fused to a human constant region.

"Humanized" anti-NPC1L1 antibodies are also within the scope of the present invention. Humanized forms of non-human (*e.g.*, murine) antibodies are chimeric immunoglobulins, which contain minimal sequence derived from non-

human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a complementary determining region of the recipient are replaced by residues from a complementary determining region of a nonhuman species (donor antibody), such as mouse, rat or rabbit, having a desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are also replaced by corresponding non-human residues.

“Single-chain Fv” or “sFv” antibody fragments include the V_H and/or V_L domains of an antibody, wherein these domains are present in a single polypeptide chain. Generally, the sFv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. Techniques described for the production of single chain antibodies (U.S. Patent Nos. 5,476,786; 5,132,405 and 4,946,778) can be adapted to produce anti-NPC1L1 specific, single chain antibodies. For a review of sFv see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore, eds., Springer-Verlag, N.Y., pp. 269-315 (1994).

“Disulfide stabilized Fv fragments” and “dsFv” include molecules having a variable heavy chain (V_H) and/or a variable light chain (V_L) which are linked by a disulfide bridge.

Antibody fragments within the scope of the present invention also include F(ab)₂ fragments which may be produced by enzymatic cleavage of an IgG by, for example, pepsin. Fab fragments may be produced by, for example, reduction of F(ab)₂ with dithiothreitol or mercaptoethylamine.

An FV fragment is a V_L or V_H region.

Depending on the amino acid sequences of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are at least five major classes of immunoglobulins: IgA, IgD, IgE, IgG and IgM, and several of these may be further divided into subclasses (isotypes), *e.g.*, IgG-1, IgG-2, IgG-3 and IgG-4; IgA-1 and IgA-2.

The anti-NPC1L1 antibody molecules of the invention may also be conjugated to a chemical moiety. The chemical moiety may be, *inter alia*, a polymer, a radionuclide or a cytotoxic factor. Preferably, the chemical moiety is a polymer which increases the half-life of the antibody molecule in the body of a subject.

Suitable polymers include, but are by no means limited to, polyethylene glycol (PEG) (e.g., PEG with a molecular weight of 2kDa, 5kDa, 10kDa, 12kDa, 20kDa, 30kDa or 40kDa), dextran and monomethoxypolyethylene glycol (mPEG). Methods for producing PEGylated anti-IL8 antibodies which are described in U.S. Patent No.

5 6,133,426 can be applied to the production of PEGylated anti-NPC1L1 antibodies of the invention. Lee, *et al.*, (1999) (Bioconj. Chem. 10: 973-981) discloses PEG conjugated single-chain antibodies. Wen, *et al.*, (2001) (Bioconj. Chem. 12: 545-553) discloses conjugating antibodies with PEG which is attached to a radiometal chelator (diethylenetriaminopentaacetic acid (DTPA)).

10 The antibody molecules of the invention may also be conjugated with labels such as ⁹⁹Tc, ⁹⁰Y, ¹¹¹In, ³²P, ¹⁴C, ¹²⁵I, ³H, ¹³¹I, ¹¹C, ¹⁵O, ¹³N, ¹⁸F, ³⁵S, ⁵¹Cr, ⁵⁷To, ²²⁶Ra, ⁶⁰Co, ⁵⁹Fe, ⁵⁷Se, ¹⁵²Eu, ⁶⁷CU, ²¹⁷Ci, ²¹¹At, ²¹²Pb, ⁴⁷Sc, ¹⁰⁹Pd, ²³⁴Th, ⁴⁰K, ¹⁵⁷Gd, ⁵⁵Mn, ⁵²Tr or ⁵⁶Fe.

The antibody molecules of the invention may also be conjugated with
15 fluorescent or chemilluminescent labels, including fluorophores such as rare earth chelates, fluorescein and its derivatives, rhodamine and its derivatives, isothiocyanate, phycoerythrin, phycocyanin, allophycocyanin, o-phthalaldehyde, fluorescamine, ¹⁵²Eu, dansyl, umbelliferone, luciferin, luminal label, isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester
20 label, an aequorin label, 2,3-dihydrophthalazinediones, biotin/avidin, spin labels and stable free radicals.

The antibody molecules may also be conjugated to a cytotoxic factor such as diptheria toxin, *Pseudomonas aeruginosa* exotoxin A chain, ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins and
25 compounds (e.g., fatty acids), dianthin proteins, *Phytolacca americana* proteins PAPI, PAPII, and PAP-S, *momordica charantia* inhibitor, curcin, crotin, *saponaria officinalis* inhibitor, mitogellin, restrictocin, phenomycin, and enomycin.

Any method known in the art for conjugating the antibody molecules of the invention to the various moieties may be employed, including those methods
30 described by Hunter, *et al.*, (1962) Nature 144: 945; David, *et al.*, (1974) Biochemistry 13: 1014; Pain, *et al.*, (1981) J. Immunol. Meth. 40: 219; and Nygren, J., (1982) Histochem. and Cytochem. 30: 407.

Methods for conjugating antibodies are conventional and very well known in the art.

Screening Assays

The invention allows the identification of selective ligands of NPC1L1 (e.g., SEQ ID NO: 2, 4 or 12) that may be useful in treatment and management of a variety of medical conditions, including elevated serum sterol (e.g., cholesterol) or 5 α -stanol. Thus, NPC1L1 of this invention can be employed in screening systems to identify ligands. These ligands may be agonists or antagonists of NPC1L1. Essentially, these assays provide methods for identifying ligands of NPC1L1 by using (1) NPC1L1, (2) an appropriate known NPC1L1 ligand, agonist or antagonist, for example, a sterol (such as cholesterol, phytosterols, including, but not limited to, sitosterol, campesterol, stigmasterol and avenosterol), a cholesterol oxidation product, a 5 α -stanol (including, but not limited to, cholestanol, 5 α -campestanol and 5 α -sitostanol), a substituted azetidinone (e.g., ezetimibe), BODIPY-ezetimibe (Altmann, *et al.*, (2002) *Biochim. Biophys. Acta* 1580(1): 77-93) or 4'', 6''-bis[(2-fluorophenyl)carbamoyl]-beta-D-cellobiosyl derivative of 11-ketotigogenin as described in DeNinno, *et al.*, (1997) (*J. Med. Chem.* 40(16): 2547-54) or any substituted azetidinone, and (3) a sample to be tested for the presence of a candidate NPC1L1 ligand.

The term "specific" when used to describe binding of, for example, a ligand of NPC1L1 in a screening assay is a term of art which refers to the extent by which the ligand or antagonist (e.g., substituted azetidinone, ezetimibe, sterol (such as cholesterol) or 5 α -stanol) binds preferentially to NPC1L1 in comparison to other proteins in the assay system. For example, detection of the specific binding of a ligand of NPC1L1 binds specifically to NPC1L1 is made apparent when a signal generated in the assay to indicate such binding exceeds, to any extent, a signal generated in a negative control wherein, for example, NPC1L1 or ligand is absent. Furthermore, "specific binding" includes binding of a ligand either directly to NPC1L1 or indirectly, for example via another moiety, in a complex of which NPC1L1 is a part. The moiety to which an NPC1L1 ligand binds can be another protein or a post-translational modification of NPC1L1 (e.g., a lipid chain or a carbohydrate chain).

Non-limiting examples of suitable substituted azetidinones for use in the screening assays include those disclosed in U.S. Patent Nos. RE37,721; 5,631,365; 5,767,115; 5,846,966; 5,688,990; 5,656,624; 5,624,920; 5,698,548; 5,756,470; 5,688,787; 5,306,817; 5,633,246; 5,627,176; 5,688,785; 5,744,467; 5 5,846,966; 5,728,827; 6,632,933 and U.S. Patent Application Publication No 2003/0105028-each of which is herein incorporated by reference in its entirety.

The present invention provides for a method by which to evaluate whether a sample contains an NPC1L1 ligand by determining whether the sample contains a candidate compound which competes for binding between the known 10 ligand (*e.g.*, ezetimibe, ezetimibe-glucuronide, compound 2, etc.) and NPC1L1. The ligand may be an agonist or antagonist. In an embodiment of the invention, the binding of the known ligand (*e.g.*, ezetimibe, ezetimibe-glucuronide, compound 2, etc.) to NPC1L1 is disrupted. The term "known ligand" refers to a compound which is known to bind to NPC1L1 and which can be detectably labeled for use in the 15 screening assays and methods described herein. "Known ligands" include the substituted 2-azetidinone glucuronides which can be detectably labeled for use in screening assays as described herein.

Ezetimibe can be prepared by a variety of methods well known to those skilled in the art, for example such as are disclosed in U.S. Patents Nos. 5,631,365, 20 5,767,115, 5,846,966, 6,207,822, U.S. Patent Application Publication No. 2002/0193607 and PCT Patent Application WO 93/02048, each of which is incorporated herein by reference in its entirety.

"Sample", "candidate compound" or "candidate substance" refers to a compound or composition which is evaluated in a test or assay, for example, for the 25 ability to bind to NPC1L1 (*e.g.*, SEQ ID NO: 2, 4 or 12) or a functional fragment thereof. The composition may comprise candidate compounds, such as small molecules, peptides, nucleotides, polynucleotides, subatomic particles (*e.g.*, α particles, β particles) or antibodies.

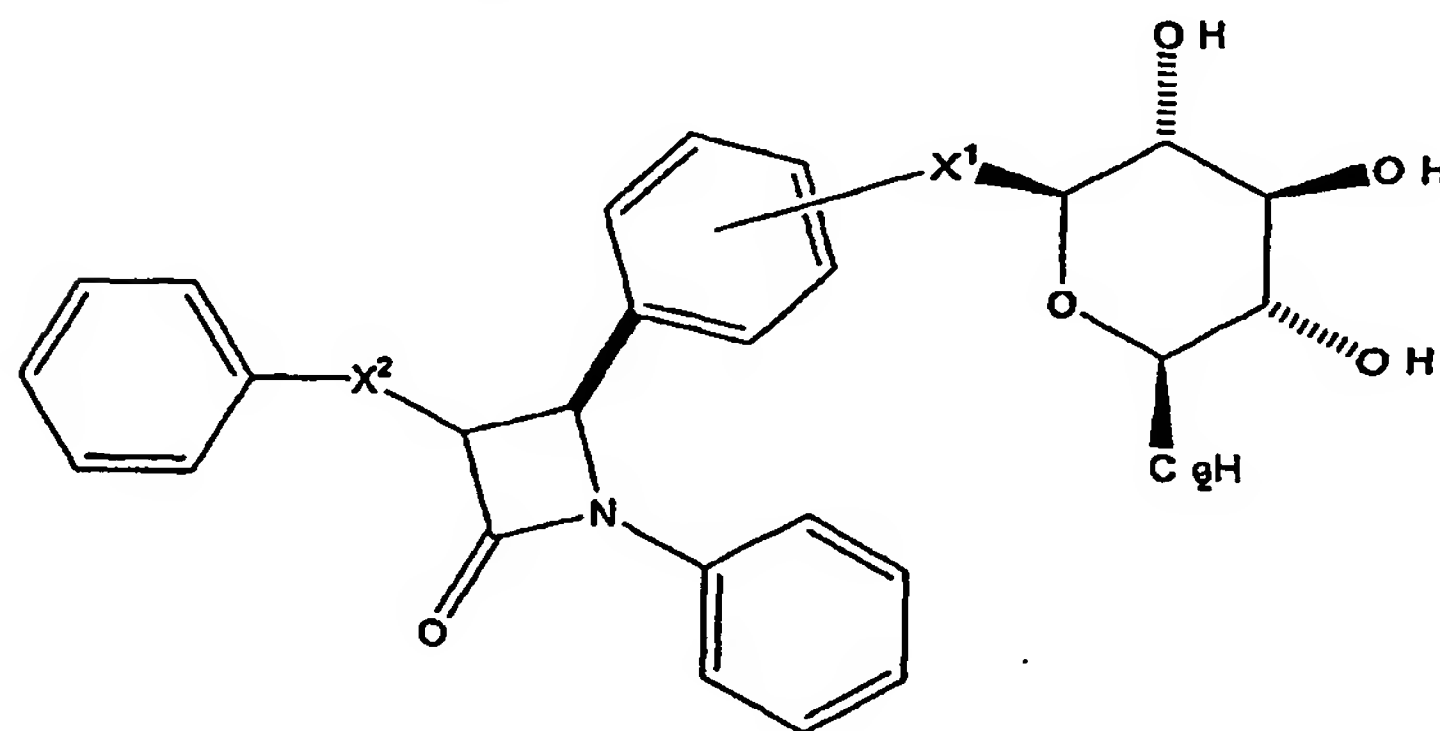
The present invention provides methods for identifying ligands of a 30 compound that binds to NPC1L1 which involve contacting NPC1L1 with a detectably labeled substituted 2-azetidinone, preferably substituted 2-azetidinone-glucuronide, and a candidate compound, and determining whether the candidate compound binds to NPC1L1, wherein binding of said candidate compound to NPC1L1 modulates

binding of the detectably labeled substituted 2-azetidinone to NPC1L1. The modulation of the binding of the substituted 2-azetidinone to NPC1L1 by the binding of the candidate compound to NPC1L1 indicates that the candidate compound is a ligand that binds to NPC1L1. It is also a good indication that the candidate compound may be an inhibitor of sterol and 5α -stanol absorption in vivo.

The present invention also provides a method for identifying a ligand of NPC1L1 comprising contacting NPC1L1 with a detectably labeled substituted 2-azetidinone, preferably substituted 2-azetidinone-glucuronide, and measuring the binding of NPC1L1 of the detectably labeled substituted 2-azetidinone in the presence and absence of a candidate compound, wherein decreased binding of the detectably labeled substituted 2-azetidinone to the NPC1L1 in the presence of the candidate compound indicates that said candidate compound is a ligand of NPC1L1 and is an inhibitor of sterol and 5α -stanol absorption.

The substituted 2-azetidinone is detectably labeled with ^3H , ^{35}S , ^{125}I , or a fluorescently labeled substituted 2-azetidinone. Preferably, the substituted 2-azetidinone is labeled with ^{35}S or ^{125}I , and particularly ^{35}S .

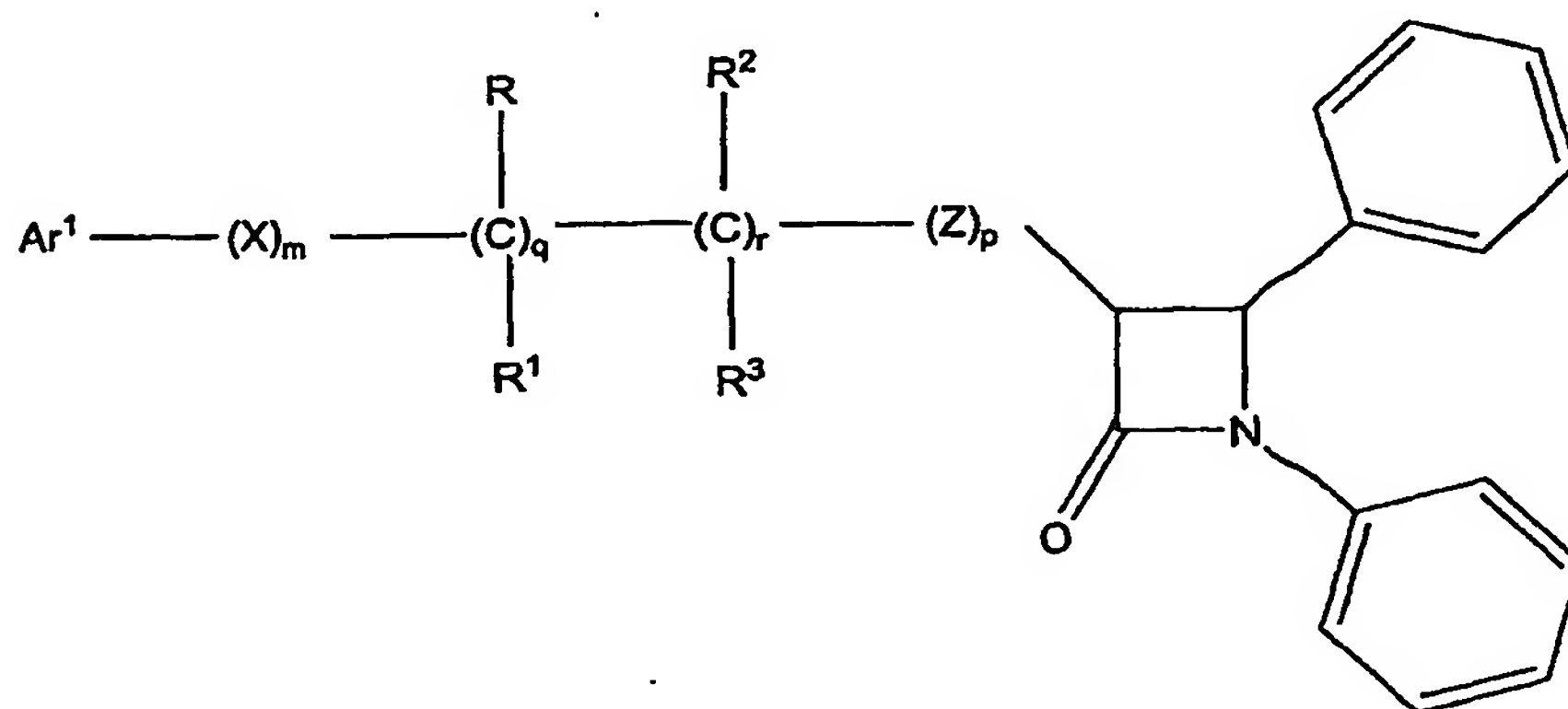
Preferably, the substituted 2-azetidinone is substituted 2-azetidinone-glucuronide. Compounds that are substituted 2-azetidinone-glucuronides are those having the following structure (I):



(I)

wherein X^1 represents a group that links the glucuronide to the 4-phenyl ring, for example but not limited to $-\text{O}-$ or $-\text{C}_{1-3}$ alkyl-, X^2 represents an optionally substituted $-\text{alkanediy}-$, and wherein any of the phenyl groups may be optionally substituted. Examples of the phenyl- X^2 -moiety in structure (I) include

those represented at the 4-position on the 2-azetidinone structure shown below in structure (II). Additional examples of substituted 2-azetidinone-glucuronides include but are not limited to those described in U.S. Patent No. 5,756,470, WO02/066464 and US 2002/0137689. Additional examples of substituted 2-azetidinone-glucuronide compounds include those having the structure (II) and pharmaceutically acceptable salts and esters thereof as follows:



10

(II)

wherein:

Ar¹ is selected from the group consisting of aryl and R⁴-substituted aryl;

X, Y and Z are independently selected from the group consisting of -CH₂-, -CH(C₁-6alkyl)- and -C(C₁-6alkyl)₂-;

R is selected from the group consisting of -OR⁶, -O(CO)R⁶, -O(CO)OR⁹, -O(CO)NR⁶R⁷, a sugar residue, a disugar residue, a trisugar residue and a tetrasugar residue;

R¹ is selected from the group consisting of -H, -C₁-6alkyl and aryl, or R and R¹ together are oxo;

R² is selected from the group consisting of -OR⁶, -O(CO)R⁶, -O(CO)OR⁹ and -O(CO)NR⁶R⁷;

R³ is selected from the group consisting of -H, -C₁-6alkyl and aryl or R² and R³ together are oxo;

q, r and t are each independently selected from 0 and 1;

m, n and p are each independently selected from 0, 1, 2, 3 and 4;

R⁴ is 1-5 substituents independently selected at each occurrence from the group consisting of:

-OR⁵, -O(CO)R⁵, -O(CO)OR⁸, -O-C₁₋₅alkyl-OR⁵, -O(CO)NR⁵R⁶, -NR⁵R⁶, -NR⁵(CO)R⁶, -NR⁵(CO)OR⁸, -NR⁵(CO)NR⁶R⁷, -NR⁵SO₂R⁸, -COOR⁵, -CONR⁵R⁶, -COR⁵, -SO₂NR⁵R⁶, -S(O)_tR⁸, -O-C₁₋₁₀alkyl-COOR⁵, -O-C₁₋₁₀alkyl-CONR⁵R⁶ and fluoro;

R⁵, R⁶ and R⁷ are independently selected at each occurrence from the group consisting of -H, C₁₋₆alkyl, aryl and aryl-substituted C₁₋₆alkyl;

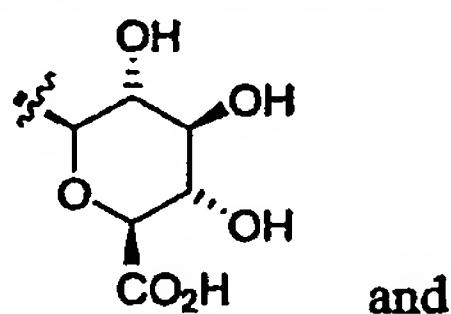
R⁸ is independently selected from the group consisting of C₁₋₆alkyl, aryl and aryl-substituted C₁₋₆alkyl;

R⁹ is selected from the group consisting of -C≡C-CH₂-NR¹⁰R¹¹, -C≡C-C(O)R¹³, and -(CH₂)₃-NR¹⁰R¹⁴;

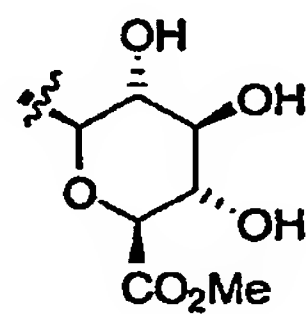
R¹⁰ is independently selected at each occurrence from -H and -C₁₋₃alkyl;

R¹¹ is selected from the group consisting of -H, -C₁₋₃alkyl, -C(O)-C₁₋₃alkyl, -C(O)-NR¹⁰R¹⁰, -SO₂-C₁₋₃alkyl, and -SO₂-phenyl; and

R¹² is selected from



and



(referred to herein as "glucuronide")

(referred to herein as "methyl ester

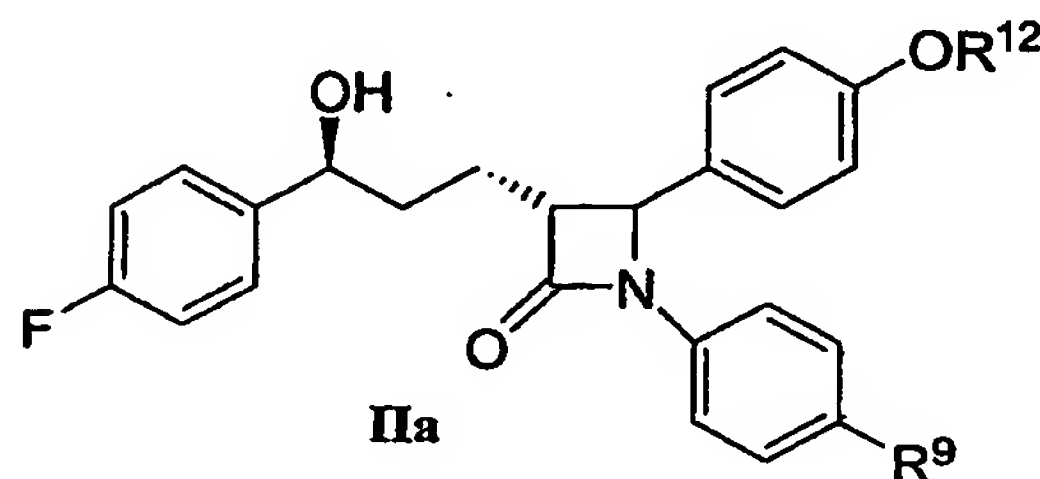
glucuronide");

R¹³ is selected from the group consisting of -OH and -NR¹⁰R¹¹; and

R¹⁴ is selected from the group consisting of -C(O)-C₁₋₃alkyl, -C(O)-NR¹⁰R¹⁰, -SO₂-C₁₋₃alkyl and -SO₂-phenyl.

In one embodiment of Formula II are compounds wherein q, r and t are each independently selected from 0 and 1; and m, n and p are each independently selected from 0, 1, 2, 3 and 4; provided that at least one of q and r is 1, and the sum of

m, n, p, q are r is 1, 2, 3, 4, 5 or 6; and provided that when p is 0 and r is 1, the sum of m, q and n is 1, 2, 3, 4, or 5. In a second embodiment of Formula II are compounds of Formula IIa,



- 5 In a class of each of these embodiments are compounds wherein R⁹ is -C≡C-CH₂-NR¹⁰R¹¹. In another class of each of these embodiments are compounds wherein R⁹ contains an -SO₂- group, i.e., wherein R⁹ is selected from the group consisting of -C≡C-CH₂-NR¹⁰R¹¹, -C≡C-C(O)NR¹⁰R¹¹, -(CH₂)₃-NR¹⁰-SO₂-C₁₋₃alkyl and -(CH₂)₃-NR¹⁰-SO₂-phenyl, and R¹¹ is selected from -SO₂-C₁₋₃alkyl, and -SO₂-phenyl.

The term "alkyl" is intended to include both branched- and straight-chain saturated aliphatic univalent hydrocarbon groups having the specified number of carbon atoms. Examples of alkyl groups include, but are not limited to, methyl (Me), ethyl (Et), n-propyl (Pr), n-butyl (Bu), n-pentyl, n-hexyl, and the isomers thereof such as isopropyl (i-Pr), isobutyl (i-Bu), secbutyl (s-Bu), tertbutyl (t-Bu), isopentyl, isohexyl and the like. If there is no specified prefix (such as "n-" for normal, "s-" for sec, "t-" for tert, "i-" for iso) with a named alkyl group, then it is intended that the named alkyl group is an n-alkyl group (i.e., "propyl" is "n-propyl"). The term "aryl" is intended to include phenyl (Ph), naphthyl, indenyl, tetrahydronaphthyl or indanyl. Phenyl is preferred.

Suitable protecting groups (designated as "PG" herein) for the hydroxyl groups of R¹² when R¹² is a glucuronide or methyl ester glucuronide include but are not limited to those that are known to be useful as carbohydrate protecting groups, such as for example benzyl, acetyl, benzoyl, *tert*-butyldiphenylsilyl, trimethylsilyl, *para*-methoxybenzyl, benzylidene, and methoxy methyl. Conditions required to selectively add and remove such protecting groups are

found in standard textbooks such as Greene, T, and Wuts, P. G. M., *Protective Groups in Organic Synthesis*, John Wiley & Sons, Inc., New York, NY, 1999.

Compounds of Formula II may contain one or more asymmetric centers and can thus occur as racemates and racemic mixtures, single enantiomers, enantiomeric mixtures, diastereomeric mixtures and individual diastereomers, and all such isomeric forms are within the scope of Formula II.

Radioactive isotopes of the compounds of Formula II are particularly useful in such assays, for example compounds of Formula II wherein sulfur is replaced with "hot" ^{35}S -, and particularly wherein the radioactive sulfur isotope is incorporated within the R^9 moiety. The use of all such radioactive isotopes of the compounds of Formula II in an assay for identifying NPC1L1 ligands is included within the scope of this invention.

The term "pharmaceutically acceptable salts" means non-toxic salts of the compounds of Formula II which are generally prepared by reacting the free acid with a suitable organic or inorganic base, particularly those formed from cations such as sodium, potassium, aluminum, calcium, lithium, magnesium, zinc and tetramethylammonium, as well as those salts formed from amines such as ammonia, ethylenediamine, N-methylglucamine, lysine, arginine, ornithine, choline, N,N'-dibenzylethylenediamine, chlorprocaine, diethanolamine, procaine, N-benzylphenethylamine, 1-p-chlorobenzyl-2-pyrrolidine-1'-yl-methylbenzimidazole, diethylamine, piperazine, morpholine, 2,4,4-trimethyl-2-pentamine and tris(hydroxymethyl)aminomethane.

When the compounds of Formula II are basic, salts may be prepared from pharmaceutically acceptable non-toxic acids, including inorganic and organic acids. Such acids include acetic, benzenesulfonic, benzoic, camphorsulfonic, citric, ethanesulfonic, fumaric, gluconic, glutamic, hydrobromic, hydrochloric, isethionic, lactic, maleic, malic, mandelic, methanesulfonic, mucic, nitric, pamoic, pantothenic, phosphoric, succinic, sulfuric, tartaric, p-toluenesulfonic acid, and the like. Particularly preferred are citric, hydrobromic, hydrochloric, maleic, phosphoric, sulfuric, and tartaric acids.

Examples of pharmaceutically acceptable esters include, but are not limited to, $-\text{C}_{1-4}$ alkyl and $-\text{C}_{1-4}$ alkyl substituted with phenyl, dimethylamino and acetylamino. " C_{1-4} alkyl" herein includes straight or branched aliphatic chains

containing from 1 to 4 carbon atoms, for example methyl, ethyl, n-propyl, n-butyl, iso-propyl, sec-butyl and tert-butyl.

The compounds of structural Formula II can be prepared according to the procedures of the following Scheme using appropriate materials, and are further exemplified by specific examples which follow. A variety of chromatographic techniques may be employed in the preparation of the compounds of Formula II. These techniques include, but are not limited to: High Performance Liquid Chromatography (including normal- reversed- and chiral-phase); Super Critical Fluid Chromatography; preparative Thin Layer Chromatography; flash chromatography with silica gel or reversed-phase silica gel; ion-exchange chromatography; and radial chromatography. All temperatures are degrees Celsius unless otherwise noted.

Some abbreviations used herein include:

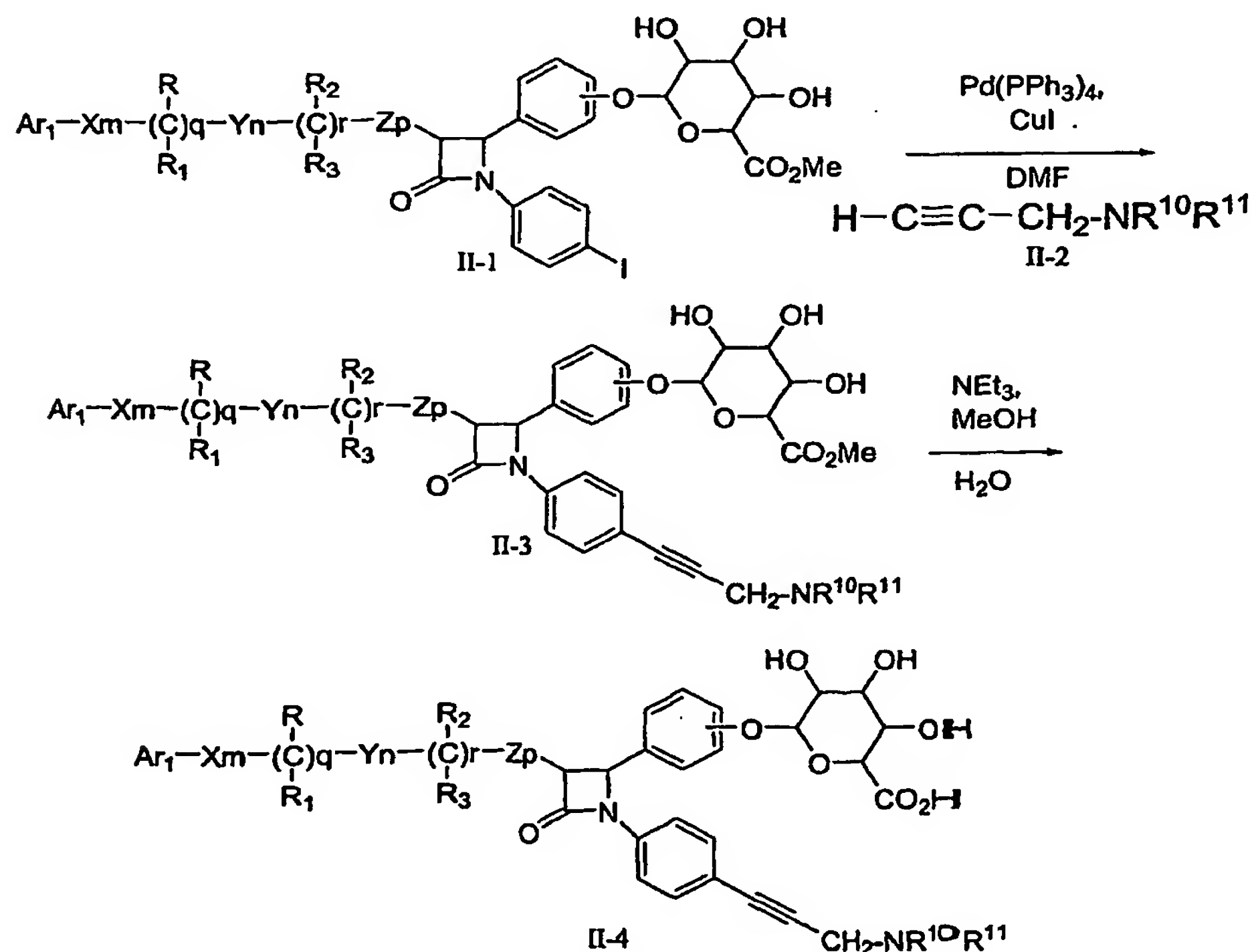
Ac	Acyl (CH ₃ C(O)-)
Bn	benzyl
calc.	Calculated
Celite	Celite™ diatomaceous earth
Dess-Martin Periodinane	1,1,1-tris(acetyloxy)-1,1-dihydro-1,2-benzodioxol-3-(1H)-one
DMF	N,N-dimethylformamide
equiv.	Equivalent(s)
ES-MS	Electron Spray Ion-Mass Spectroscopy
EtOAc	Ethyl acetate
h	Hour(s)
HPLC	High performance liquid chromatography
min	Minute(s)
m.p.	Melting point
MS	Mass spectrum
r.t. (or rt)	Room temperature
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Tlc	Thin layer chromatography

The general Scheme below illustrates a method for the syntheses of compounds of structural formula II-4. All substituents are as defined in Formula II unless indicated otherwise. In this method, II-1 is treated with a terminal alkyne of

type II-2 in the presence of a suitable palladium catalyst such as tetrakis(triphenylphosphine) palladium(0) or [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium(II) or the like, and copper(I) iodide. The reaction is usually performed in an inert organic solvent such as DMF, between room temperature and 100 °C, for a period of 6-48 h, and the product is an internal alkyne of structural formula II-3. Alkyne II-2 may contain a radioactive atom such as ³⁵S to provide the corresponding radiolabeled adduct upon reaction with II-1. Conversion of II-3 to II-4 can be achieved using a variety of hydrolytic methods known to those skilled in the art of organic synthesis. For example, a particularly mild hydrolysis protocol involves the treatment of II-3 with a tertiary amine base such as triethylamine, or diisopropylethylamine or the like, in a mixed solvent system comprising methanol and water. The product of the reaction is a compound of structural formula II-4. By utilizing the procedures described herein, one of ordinary skill in the art can readily prepare additional compounds of Formula II.

15

SCHEME



Two additional types of screening systems that can be used include a labeled-ligand binding assay (*e.g.*, direct binding assay or scintillation proximity assay (SPA)) and a “sterol (*e.g.*, cholesterol) or 5 α -stanol uptake” assay. A labeled ligand, for use in the binding assay, can be obtained by labeling a sterol (*e.g.*, cholesterol) or a 5 α -stanol or a known NPC1L1 agonist or antagonist with a measurable group (*e.g.*, ^{35}S , ^{125}I or ^3H). Various labeled forms of sterols (*e.g.*, cholesterol) or 5 α -stanols are available commercially or can be generated using standard techniques (*e.g.*, Cholesterol- [1,2- $^3\text{H}(\text{N})$], Cholesterol-[1,2,6,7- $^3\text{H}(\text{N})$] or Cholesterol-[7- $^3\text{H}(\text{N})$]; American Radiolabeled Chemicals, Inc; St. Louis, MO). In a preferred embodiment, ezetimibe is fluorescently labeled with a BODIPY group (Altmann, *et al.*, (2002) *Biochim. Biophys. Acta* 1580(1): 77-93) or labeled with a detectable group such as ^{35}S , ^{125}I or ^3H , preferably ^{35}S .

Direct Binding Assay. Typically, a given amount of NPC1L1 of the invention (*e.g.*, SEQ ID NO: 2, 4 or 12) or a complex including NPC1L1 is contacted with increasing amounts of labeled ligand or known antagonist or agonist (discussed above) and the amount of the bound, labeled ligand or known antagonist or agonist is measured after removing unbound, labeled ligand or known antagonist or agonist by washing. As the amount of the labeled ligand or known agonist or antagonist is increased, a point is eventually reached at which all receptor binding sites are occupied or saturated. Specific receptor binding of the labeled ligand or known agonist or antagonist is abolished by a large excess of unlabeled ligand or known agonist or antagonist.

Preferably, an assay system is used in which non-specific binding of the labeled ligand or known antagonist or agonist to the receptor is minimal. Non-specific binding is typically less than 50%, preferably less than 15%, more preferably less than 10%, and most preferably 5% or less, of the total binding of the labeled ligand or known antagonist or agonist.

In the basic binding assay, the method for identifying an NPC1L1 ligand, agonist or antagonist includes:

- (a) contacting NPC1L1 (*e.g.*, SEQ ID NO: 2 or 4 or 12), a fragment thereof or a complex including NPC1L1, in the presence of a known amount of labeled sterol (*e.g.*, cholesterol) or 5 α -stanol or known antagonist or agonist (*e.g.*,

labeled ezetimibe) with a sample to be tested for the presence of an NPC1L1 ligand, agonist or antagonist; and

(b) measuring the amount of labeled sterol (e.g., cholesterol) or 5 α -stanol or known antagonist or agonist directly or indirectly bound to NPC1L1.

5 An NPC1L1 ligand in the sample is identified by measuring substantially reduced direct or indirect binding of the labeled sterol (e.g., cholesterol) or 5 α -stanol or known antagonist or agonist to NPC1L1, compared to what would be measured in the absence of such a ligand. For example, reduced direct or indirect binding between [³H]-cholesterol and NPC1L1 in the presence of a sample might
10 suggest that the sample contains a substance which is competing against [³H]-cholesterol for NPC1L1 binding.

This assay can include a control experiment lacking any NPC1L1-dependent ligand (e.g., sterol such as cholesterol or 5 α -stanol) binding. In this assay, for example, a whole cell or cell membrane lacking any functional NPC1L1, for
15 example, a cell or membrane isolated or derived from a transgenic mutant *npc1l1*^{-/-} mouse of the invention, is assayed for ligand binding. When screening a sample for the presence of an NPC1L1 antagonist, it is useful to compare the level of binding observed in the presence of a sample being tested with that of a control experiment, as described herein, which completely lacks NPC1L1-dependent binding. Ideally,
20 though by no means necessarily, the level of binding seen in the presence of a sample containing an antagonist will be similar to that of the control experiment.

Alternatively, a sample can be tested directly for binding to NPC1L1 (e.g., SEQ ID NO: 2, 4 or 12). A basic assay of this type may include the following steps:

25 (a) contacting NPC1L1 (e.g., SEQ ID NO: 2 or 4 or 12), a fragment thereof or a complex including NPC1L1 with a labeled candidate compound (e.g., [³H]-ezetimibe); and

(b) detecting direct or indirect binding between the labeled candidate compound and NPC1L1.

30 Again, these experiment can be performed along with a control experiment wherein NPC1L1-dependent binding is completely lacking. For example, the assay can be performed using a whole cell or cell membrane lacking any

functional NPC1L1 (e.g., cell or cell membrane derived from a transgenic, mutant *npc1l1* mouse as described herein).

A candidate compound which is found to bind to NPC1L1 may function as ligand, agonist or antagonist of NPC1L1 (e.g., by inhibition of sterol
5 (e.g., cholesterol) or 5 α -stanol uptake).

In an embodiment of the invention, the bound candidate compound is quantified after filtration using glass fiber filters. In one aspect of this embodiment, the bound candidate compound is detected after single-tube vacuum filtration of GF/C glass fiber filters, obtained from Whatman. The filters may be pretreated by soaking
10 with 0.5% polyethylenimine to reduce nonspecific binding. Filtration is accomplished by adding ice cold buffer to the assay tube, pouring the mixture through the filter, and then rinsing the tube and filter twice more with additional buffer. The buffer may be a Tris buffer or MES buffer (120 mM NaCl, 0.1% sodium cholate, and 20 mM MES at pH 6.70). The filters can be counted using scintillation fluid, e.g., Packard DM
15 liquid or Packard Ultima Gold MV.

Alternatively, vacuum filtration of the sample on a Millipore 96-well plate (Whatman GF/C) can also be used to achieve adequate precision in a manner well-known to those skilled in the art.

SPA Assay. NPC1L1 ligands may also be measured using scintillation
20 proximity assays (SPA). SPA assays are conventional and very well known in the art; see, for example, U.S. Patent No. 4,568,649. In SPA, the target of interest is immobilized to a small microsphere approximately 5 microns in diameter. The microsphere, typically, includes a solid scintillant core which has been coated with a polyhydroxy film, which in turn contains coupling molecules, which allow generic
25 links for assay design. When a radioisotopically labeled molecule binds to the microsphere, the radioisotope is brought into close proximity to the scintillant and effective energy transfer from electrons emitted by the isotope will take place resulting in the emission of light. While the radioisotope remains in free solution, it is too distant from the scintillant and the electron will dissipate the energy into the
30 aqueous medium and therefore remain undetected. Scintillation may be detected with a scintillation counter. In general, ^3H , ^{125}I and ^{35}S labels are well suited to SPA.

For the assay of receptor-mediated binding events, the lectin wheat germ agglutinin (WGA) may be used as the SPA bead coupling molecule (Amersham

Biosciences; Piscataway, NJ). The WGA coupled bead captures glycosylated, cellular membranes and glycoproteins and has been used for a wide variety of receptor sources and cultured cell membranes. The receptor is immobilized onto the WGA-SPA bead and a signal is generated on binding of an isotopically labeled
5 ligand. Other coupling molecules which may be useful for receptor binding SPA assays include poly-L-lysine and WGA/polyethyleneimine (Amersham Biosciences; Piscataway, NJ). See, for example, Berry, J.A., *et al.*, (1991) Cardiovascular Pharmacol. 17 (Suppl.7): S143-S145; Hoffman, R., *et al.*, (1992) Anal. Biochem. 203: 70-75; Kienhus, *et al.*, (1992) J. Receptor Research 12: 389-399; Jing, S., *et al.*,
10 (1992) Neuron 9: 1067-1079.

The scintillant contained in SPA beads may include, for example, yttrium silicate (YSi), yttrium oxide (YOx), diphenyloxazole or polyvinyltoluene (PVT) which acts as a solid solvent for diphenylanthracene (DPA).

SPA assays may be used to analyze whether a sample contains an
15 NPC1L1 ligand. In these assays, a host cell which expresses NPC1L1 (*e.g.*, SEQ ID NO: 2 or 4 or 12) on the cell surface or a membrane fraction thereof is incubated with and captured by SPA beads (*e.g.*, WGA coated YOx beads or WGA coated YSi beads). The beads bearing the NPC1L1 are incubated with labeled, known ligand or agonist or antagonist (*e.g.*, ³H-cholesterol, ³H-ezetimibe, ¹²⁵I-ezetimibe or a ³⁵S-
20 ezetimibe analog). The assay mixture further includes either the sample to be tested or a blank (*e.g.*, water). After an optional incubation, scintillation is measured using a scintillation counter. An NPC1L1 ligand, agonist or antagonist may be identified in the sample by measuring substantially reduced fluorescence, compared to what would be measured in the absence of such ligand, agonist or antagonist (blank). Measuring
25 substantially reduced fluorescence may suggest that the sample contains a substance which competes for direct or indirect NPC1L1 binding with the known ligand, agonist or antagonist.

Alternatively, a sample may be identified as an ligand of NPC1L1 by directly detecting binding in a SPA assay. In this assay, a labeled version of a
30 candidate compound to be tested may be put in contact with the host cell expressing NPC1L1 or a membrane fraction thereof which is bound to the SPA bead. Fluorescence may then be assayed to detect the presence of a complex between the labeled candidate compound and the host cell or membrane fraction expressing

NPC1L1 or a complex including NPC1L1. A candidate compound which binds directly or indirectly to NPC1L1 may possess NPC1L1 agonistic or antagonistic activity.

SPA Assays can also be performed along with a control experiment lacking any NPC1L1-dependent binding. The control experiment can be performed, for example, with a cell or cell membrane lacking any functional NPC1L1 (e.g., cell or cell membrane derived from a transgenic, mutant *npc1l1*- mouse as described herein). When the control experiment is performed, the level of binding observed in the presence of sample being tested for the presence of an antagonist can be compared with that observed in the control experiment.

Sterol/5 α -stanol Uptake Assay. Assays may also be performed to determine if a sample can agonize or antagonize NPC1L1 mediated sterol (e.g., cholesterol) or 5 α -stanol uptake. In these assays, a host cell expressing NPC1L1 (e.g., SEQ ID NO: 2 or 4 or 12) on the cell surface (discussed above) can be contacted with detectably labeled sterol (e.g., ³H-cholesterol or ¹²⁵I-cholesterol) or 5 α -stanol along with either a sample or a blank. After an optional incubation, the cells can be washed to remove unabsorbed sterol or 5 α -stanol. Sterol or 5 α -stanol uptake can be determined by detecting the presence of labeled sterol or 5 α -stanol in the host cells. For example, assayed cells or lysates or fractions thereof (e.g., fractions resolved by thin-layer chromatography) can be contacted with a liquid scintillant and scintillation can be measured using a scintillation counter.

In these assays, an NPC1L1 antagonist in the sample may be identified by measuring substantially reduced uptake of labeled sterol (e.g., ³H-cholesterol) or 5 α -stanol, compared to what would be measured in the absence of such an antagonist and an agonist may be identified by measuring substantially increased uptake of labeled sterol (e.g., ³H-cholesterol) or 5 α -stanol, compared to what would be measured in the absence of such an agonist.

Uptake assays can also be performed along with a control experiment lacking any NPC1L1-dependent uptake. The control experiment can be performed, for example, with a cell lacking any functional NPC1L1 (e.g., cell derived from a transgenic, mutant *npc1l1*- mouse as described herein). When the control experiment is performed, the level of uptake observed in the presence of sample being tested for

the presence of an antagonist can be compared with that observed in the control experiment.

Source of NPC1L1. In principle, a binding assay of the invention could be carried out using a soluble NPC1L1 polypeptide of the invention, *e.g.*, following production and refolding by standard methods from an *E. coli* or other prokaryotic or eukaryotic expression system, and the resulting receptor-labeled ligand complex could be precipitated, *e.g.*, using an antibody against the receptor. The precipitate could then be washed and the amount of the bound, labeled ligand or antagonist or agonist could be measured.

Alternatively, NPC1L1 is membrane-bound. A nucleic acid encoding an NPC1L1 polypeptide of the invention (*e.g.*, SEQ ID NO: 2, 4 or 12) can be transfected into an appropriate host cell, whereby the NPC1L1 will become incorporated into the membrane of the cell. A membrane fraction can then be isolated from the cell and used as a source of NPC1L1 for assay. Alternatively, the whole cell expressing NPC1L1 in the cell surface can be used in an assay. Preferably, specific binding of the labeled ligand or known antagonist or agonist to an untransfected/untransformed host cell or to a membrane fraction from an untransfected/untransformed host cell will be negligible.

Various membranes may be used directly as a source of NPC1L1 for the above-described screening systems, *e.g.* direct binding, scintillation proximity assay, sterol/5 α stanol uptake assay. As described in Examples 5, 6, 7, 8, 9, 17, 27, and 29, NPC1L1 is highly expressed in certain tissues, especially in brush border cells of intestinal tissues. Therefore, brush border membrane (BBM) vesicle preparations may be utilized as a source of NPC1L1. The membranes may be derived from mammalian intestinal tissue from rhesus, rat, mouse or human tissue.

Membranes may be derived from brush border cells of intestinal tissues. Such membranes are conventionally prepared by collecting intestinal tissue from freshly sacrificed animals. The mucosa of the tissue is scraped, collected into buffered solutions, and homogenized. Cellular debris is removed and the membrane fractions are collected by centrifugation. Conventional techniques known to one of skill in the art may be used for the preparation of brush border membrane vesicles. See Hauser, H., Howell, K., Dawson, R.M.C., Bowyer, D. E. *Biochim. Biophys. Acta* 602, 567-577 (1980); Kramer, W., Girbig, F., Gutjahr, U., Kowalewski, S., Jouvenal,

K., Muller, G., Tripier, D., Wess, G. J. Biol. Chem. 268, 18035-18046 (1993);
Rigtrup, K.M., Ong, D.E. Biochemistry 31, 2920-2926 (1992).

The membrane preparation may be in vesicular or non-vesicular form.

Alternatively, liposomes and liposomal preparations comprising

5 NPC1L1 may also be a viable source of NPC1L1 for the screening assays of the present claimed method.

In vitro cultured cells expressing NPC1L1 may also be used. The host cells may be prepared by transforming or transfecting a nucleic acid encoding an NPC1L1 of the invention into an appropriate host cell, whereby the receptor becomes
10 incorporated into the membrane of the cell. A membrane fraction can then be isolated from the cell and used as a source of the receptor for assay. Alternatively, the whole cell expressing the receptor on the cell surface can be used in an assay. Preferably, specific binding of the labeled ligand or known antagonist or agonist to an untransfected/untransformed host cell or membrane fraction from an
15 untransfected/untransformed host cell will be negligible.

Preferred host cells include Chinese Hamster Ovary (CHO) cells, murine macrophage J774 cells, HEK-293 cells or any other macrophage cell line and human intestinal epithelial Caco2 cells.

The present invention provides for a method of identifying a ligand of
20 NPC1L1 using these membrane preparations, for example by contacting membranes comprising NPC1L1, such as brush border membrane vesicle preparations, with detectably labeled substituted azetidinone compounds which are known NPC1L1 ligands, agonists or antagonists, and a candidate compound and determining whether the candidate compound can bind to NPC1L1. The binding of the candidate
25 compound to NPC1L1 may modulate binding of the detectably labeled NPC1L1 ligands, agonists or antagonists to NPC1L1. In addition, a NPC1L1 ligand may be identified by measuring the binding of NPC1L1 with detectably labeled NPC1L1 ligands, agonists or antagonists in the presence and absence of the candidate compound wherein decreased binding of the detectably labeled NPC1L1 ligands,
30 agonists or antagonists to NPC1L1 is an indication that the candidate compound is ligand of NPC1L1.

NPC1L1 may also be obtained by solubilization of membrane fractions comprising NPC1L1. The membranes may be obtained as discussed above, *e.g.*, from mammalian tissue or *in vitro* cultured cells.

Binding Affinities of NPC1L1 Ligands. The affinity and specificity of the known ligand (e.g., detectably labeled substituted 2-azetidinone-glucuronide) for NPC1L1 are important to the identification of ligands that bind NPC1L1 in a screening assay. It is understood that the known ligand will be labeled for use in the screening assay. In an embodiment of the invention, the binding affinity of the known ligand for human NPC1L1 has a K_D value equivalent or lower than the K_D value of ezetimibe glucuronide **1** for human NPC1L1. In an aspect of this embodiment, the binding affinity of the known ligand for human NPC1L1 has a K_D value of about 200nM or lower; particularly it has a K_D value of about 100nM or lower; more particularly it has a K_D value of about 50nM or lower; even more particularly it has a K_D value of about 20nM or lower; and most particularly it has a K_D value of about 10nM or lower. For usefulness in the assay, there is essentially no lower limit on the K_D value of the known ligand and it may, for example, go down into the pM range. As the K_D value decreases, the binding affinity of the ligand for human NPC1L1 increases, which is desirable for the screening assay.

In another embodiment of the invention, the binding affinity of the known ligand for rat NPC1L1 has a K_D value equivalent or lower than the K_D value of ezetimibe glucuronide **1** for rat NPC1L1. In an aspect of this embodiment, the binding affinity of the known ligand for rat NPC1L1 has a K_D value of about 200nM or lower; particularly it has a K_D value of about 100nM or lower; more particularly it has a K_D value of about 50nM or lower; even more particularly it has a K_D value of about 20nM or lower; and most particularly it has a K_D value of about 10nM or lower.

In another embodiment of this invention, the known ligand for human NPC1L1 is selected from (a) a sulfur-containing substituted 2-azetidinone-glucuronide that is labeled with ^{35}S , and particularly a compound of Formula II wherein R^9 contains an $-\text{SO}_2-$ group and (b) a substituted 2-azetidinone-glucuronide labeled with ^{125}I .

In one aspect of this embodiment, the known ligand for human NPC1L1 is selected from (a) a sulfur-containing substituted 2-azetidinone-glucuronide that is labeled with ^{35}S , and particularly a compound of Formula II

wherein R⁹ contains an -SO₂- group, and (b) a substituted 2-azetidinone-glucuronide labeled with ¹²⁵I, and has a K_D value equivalent or lower than the K_D value of ezetimibe glucuronide 1.

In another aspect of this embodiment, the known ligand for human
5 NPC1L1 is selected from (a) a sulfur-containing substituted 2-azetidinone-glucuronide that is labeled with ³⁵S, and particularly a compound of Formula II wherein R⁹ contains an -SO₂- group, and (b) a substituted 2-azetidinone-glucuronide labeled with ¹²⁵I, and has a K_D value of about 200nM or lower; particularly it has a K_D value of about 100nM or lower; more particularly it has a K_D value of about 50nM or
10 lower; even more particularly it has a K_D value of about 20nM or lower; and most particularly it has a K_D value of about 10nM or lower.

When using ³H-labeled ezetimibe glucuronide in screening to identify NPC1L1 ligands from among the candidate compounds using mouse-derived membranes, candidate compounds identified as NPC1L1 ligands are preferably those
15 candidates that exhibit a binding affinity having a K_D value of about 12,000nM or lower, preferably about 1000nM or lower, more preferably about 100nM or lower, and most preferably about 10nM or lower. When using ³H-labeled ezetimibe glucuronide in screening to identify NPC1L1 ligands using rat-derived membranes or human-derived membranes, candidate compounds identified as NPC1L1 ligands are
20 preferably those candidates that exhibit a binding affinity having a K_D value of about 1000nM or lower, preferably about 100nM or lower, and more preferably about 10nM or lower. When using ³H-labeled ezetimibe glucuronide in screening to identify NPC1L1 ligands using rhesus monkey-derived membranes, candidate compounds identified as NPC1L1 ligands are preferably those candidates that exhibit a binding
25 affinity having a K_D value of about 50nM or lower, and preferably about 10nM or lower.

When using ³⁵S-labeled compound 2 in screening to identify NPC1L1 ligands from among the candidate compounds using rat or human-derived membranes, candidate compounds identified as NPC1L1 ligands are preferably those
30 candidates that exhibit a binding affinity having a K_D value in the range from about 10μM to about 1nM. When using ¹²⁵I-labeled substituted 2-azetidinone glucuronide compounds in the assay with rat or human membranes, candidate compounds identified as NPC1L1 ligands are preferably those candidates that exhibit a binding

affinity having a K_D value in the range from about 10nM to about 10pM, and preferably from about 100pM to about 10pM.

Mouse Assay. The present invention comprises a mutant, transgenic mouse which lacks any functional NPC1L1. This mouse may serve as a convenient control experiment in screening assays for identifying inhibitors of intestinal sterol (e.g., cholesterol) or 5 α -stanol absorption, preferably inhibitors of NPC1L1. Preferably, a mouse assay of the present invention would comprise the following steps:

(a) feeding a sterol (e.g., cholesterol) or 5 α -stanol-containing substance (e.g., comprising radiolabeled cholesterol, such as ^{14}C -cholesterol or ^3H -cholesterol) to a first and second mouse comprising a functional *NPC1L1* gene and to a third, mutant mouse lacking a functional NPC1L1;

The sterol (e.g., cholesterol) or 5 α -stanol containing substance preferably contains labeled cholesterol, such as a radiolabeled cholesterol, for example, ^3H or ^{14}C labeled cholesterol. The sterol (e.g., cholesterol) or 5 α -stanol containing substance may also include cold, unlabeled sterol (e.g., cholesterol) or 5 α -stanol such as in corn oil.

In these assays, the third *npc1l1*⁻ mutant mouse serves as a (+)-control experiment which exhibits low levels of intestinal sterol (e.g., cholesterol) or 5 α -stanol absorption and the second mouse serves as a (-)-control experiment which exhibits normal, uninhibited levels of intestinal sterol (e.g., cholesterol) or 5 α -stanol absorption. The second mouse is not administered the sample to be tested for an NPC1L1 antagonist. The first mouse is the experiment.

(b) administering the sample to the first mouse comprising a functional *NPC1L1* but not to the second mouse;

(c) measuring the amount of sterol (e.g., cholesterol) or 5 α -stanol absorption in the intestine of said first, second and third mouse;

Intestinal sterol (e.g., cholesterol) or 5 α -stanol absorption may be measured by any method known in the art. For example, the level intestinal absorption can be assayed by measuring the level of serum sterol (e.g., cholesterol) or 5 α -stanol.

(d) comparing the levels of intestinal sterol (e.g., cholesterol) or 5 α -stanol absorption in each mouse;

wherein the sample is determined to contain the intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption antagonist when the level of intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption in the first mouse and in the third mouse are less than the amount of intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption in the
5 second mouse.

Preferably, if the sample contains an intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption inhibitor (*e.g.*, an NPC1L1 inhibitor), the level of sterol (*e.g.*, cholesterol) or 5 α -stanol absorption in the first mouse will be similar to that of the third, *npc1l1* mutant mouse.

10 An alternative, (+)-control experiment which may be used in these screening assays is a mouse comprising functional NPC1L1 which is administered a known antagonist of NPC1L1, such as ezetimibe.

Pharmaceutical Compositions

NPC1L1 ligands discovered, for example, by the screening methods
15 described above may be used therapeutically (*e.g.*, in a pharmaceutical composition) to stimulate or block the activity of NPC1L1 and, thereby, to treat any medical condition caused or mediated by NPC1L1. In addition, the antibody molecules of the invention may also be used therapeutically (*e.g.*, in a pharmaceutical composition) to bind NPC1L1 and, thereby, block the ability of NPC1L1 to bind a sterol (*e.g.*,
20 cholesterol) or 5 α -stanol. Blocking the binding of a sterol (*e.g.*, cholesterol) or 5 α -stanol would prevent absorption of the molecule (*e.g.*, by intestinal cells such as enterocytes). Blocking absorption of sterol (*e.g.*, cholesterol) or 5 α -stanol would be a useful way to lower serum sterol (*e.g.*, cholesterol) or 5 α -stanol levels in a subject and, thereby, reduce the incidence of, for example, hyperlipidemia, atherosclerosis,
25 coronary heart disease, stroke or arteriosclerosis.

The term "subject" or "patient" includes any organism, preferably animals, more preferably mammals (*e.g.*, mice, rats, rabbits, dogs, horses, primates, cats) and most preferably humans.

The term "pharmaceutical composition" refers to a composition
30 including an active ingredient and a pharmaceutically acceptable carrier and/or adjuvant.

Although the compositions of this invention could be administered in simple solution, they are more typically used in combination with other materials such

as carriers, preferably pharmaceutically acceptable carriers. Useful, pharmaceutically acceptable carriers can be any compatible, non-toxic substances suitable for delivering the compositions of the invention to a subject. Sterile water, alcohol, fats, waxes, and inert solids may be included in a pharmaceutically acceptable carrier.

- 5 Pharmaceutically acceptable adjuvants (buffering agents, dispersing agents) may also be incorporated into the pharmaceutical composition.

Preferably, the pharmaceutical compositions of the invention are in the form of a pill or capsule. Methods for formulating pills and capsules are very well known in the art. For example, for oral administration in the form of tablets or
10 capsules, the active drug component may be combined with any oral, non-toxic pharmaceutically acceptable inert carrier, such as lactose, starch, sucrose, cellulose, magnesium stearate, dicalcium phosphate, calcium sulfate, talc, mannitol, ethyl alcohol (liquid forms) and the like. Moreover, when desired or needed, suitable binders, lubricants, disintegrating agents and coloring agents may also be incorporated
15 in the mixture. Suitable binders include starch, gelatin, natural sugars, corn sweeteners, natural and synthetic gums such as acacia, sodium alginate, carboxymethylcellulose, polyethylene glycol and waxes. Among the lubricants there may be mentioned for use in these dosage forms, boric acid, sodium benzoate, sodium acetate, sodium chloride, and the like. Disintegrants include starch, methylcellulose,
20 guar gum and the like. Sweetening and flavoring agents and preservatives may also be included where appropriate.

The pharmaceutical compositions of the invention may be administered in conjunction with a second pharmaceutical composition or substance. In preferred embodiments, the second composition includes a cholesterol-lowering
25 drug. When a combination therapy is used, both compositions may be formulated into a single composition for simultaneous delivery or formulated separately into two or more compositions (e.g., a kit).

The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well known in the art of pharmacy. See, e.g.,
30 Gilman *et al.* (eds.) (1990), *The Pharmacological Bases of Therapeutics*, 8th Ed., Pergamon Press; and Remington's *Pharmaceutical Sciences*, *supra*, Easton, Penn.; Avis *et al.* (eds.) (1993) *Pharmaceutical Dosage Forms: Parenteral Medications* Dekker, New York; Lieberman *et al.* (eds.) (1990) *Pharmaceutical Dosage Forms:*

Tablets Dekker, New York; and Lieberman *et al.* (eds.) (1990), *Pharmaceutical Dosage Forms: Disperse Systems* Dekker, New York.

The dosage regimen involved in a therapeutic application may be determined by a physician, considering various factors which may modify the action of the therapeutic substance, *e.g.*, the condition, body weight, sex and diet of the patient, the severity of any infection, time of administration, and other clinical factors. Often, treatment dosages are titrated upward from a low level to optimize safety and efficacy. Dosages may be adjusted to account for the smaller molecular sizes and possibly decreased half-lives (clearance times) following administration.

An "effective amount" of a ligand of the invention may be an amount that will detectably reduce the level of intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption or detectably reduce the level of serum sterol (*e.g.*, cholesterol) or 5 α -stanol in a subject administered the composition.

Typical protocols for the therapeutic administration of such substances are well known in the art. Pharmaceutical composition of the invention may be administered, for example, by any parenteral or non-parenteral route.

Pills and capsules of the invention can be administered orally. Injectable compositions can be administered with medical devices known in the art; for example, by injection with a hypodermic needle.

Injectable pharmaceutical compositions of the invention may also be administered with a needleless hypodermic injection device; such as the devices disclosed in U.S. Patent Nos. 5,399,163; 5,383,851; 5,312,335; 5,064,413; 4,941,880; 4,790,824 or 4,596,556.

Anti-Sense

The present invention also encompasses anti-sense oligonucleotides capable of specifically hybridizing to mRNA encoding NPC1L1 (*e.g.*, any of SEQ ID NOs: 1, 3, 5-11 or 13) having an amino acid sequence defined by, for example, SEQ ID NO: 2 or 4 or 12 or a subsequence thereof so as to prevent translation of the mRNA. Additionally, this invention contemplates anti-sense oligonucleotides capable of specifically hybridizing to the genomic DNA molecule encoding NPC1L1, for example, having an amino acid sequence defined by SEQ ID NO: 2 or 4 or 12 or a subsequence thereof.

This invention further provides pharmaceutical compositions comprising (a) an amount of an oligonucleotide effective to reduce NPC1L1-mediated sterol (*e.g.*, cholesterol) or 5 α -stanol absorption by passing through a cell membrane and binding specifically with mRNA encoding NPC1L1 in the cell so as to prevent its translation and (b) a pharmaceutically acceptable carrier capable of passing through a cell membrane. In an embodiment, the oligonucleotide is coupled to a substance that inactivates mRNA. In another embodiment, the substance that inactivates mRNA is a ribozyme.

Reducing the level of NPC1L1 expression by introducing anti-sense NPC1L1 RNA into the cells of a patient is a useful method reducing intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption and serum cholesterol in the patient.

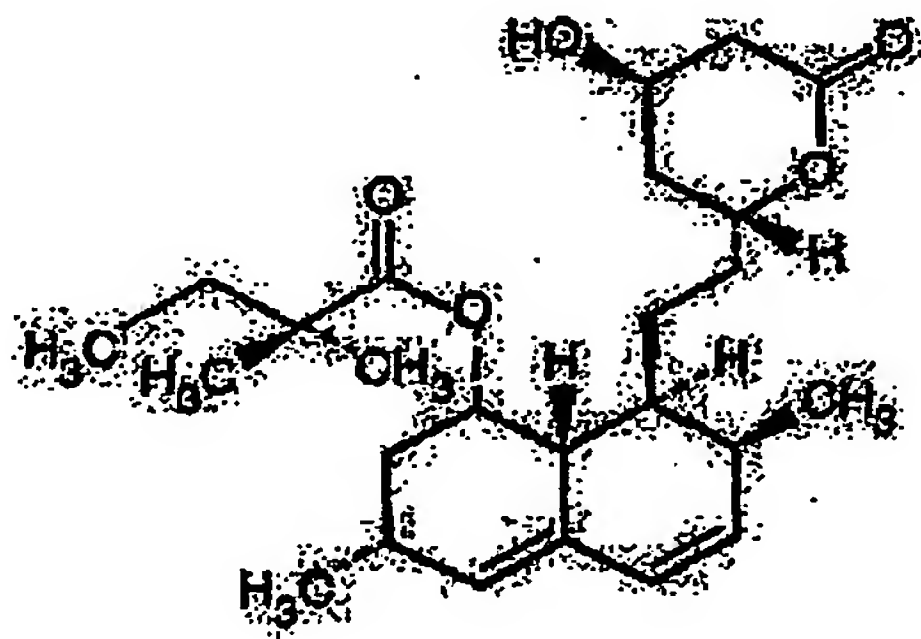
Kits

Kits of the present invention include ezetimibe, preferably combined with a pharmaceutically acceptable carrier, in a pharmaceutical formulation, more preferably in a pharmaceutical dosage form such as a pill, a powder, an injectable liquid, a tablet, dispersible granules, a capsule, a cachet or a suppository. See for example, Gilman *et al.* (eds.) (1990), *The Pharmacological Bases of Therapeutics*, 8th Ed., Pergamon Press; and Remington's *Pharmaceutical Sciences*, supra, Easton, Penn.; Avis *et al.* (eds.) (1993) *Pharmaceutical Dosage Forms: Parenteral Medications* Dekker, New York; Lieberman *et al.* (eds.) (1990) *Pharmaceutical Dosage Forms: Tablets* Dekker, New York; and Lieberman *et al.* (eds.) (1990), *Pharmaceutical Dosage Forms: Disperse Systems* Dekker, New York. Preferably, the dosage form is a Zetia® tablet (Merck/Schering-Plough Corp.). Ezetimibe may be supplied in any convenient form. For example, tablets including ezetimibe may be supplied in bottles of 30, 90 or 500.

The kits of the present invention also include information, for example in the form of a package insert, indicating that the target of ezetimibe is NPC1L1 (NPC3). The term "target of ezetimibe" indicates that ezetimibe reduces intestinal sterol (*e.g.*, cholesterol) or 5 α -stanol absorption, either directly or indirectly, by antagonizing NPC1L1. The form of the insert may take any form, such as paper or on electronic media such as a magnetically recorded medium (*e.g.*, floppy disk) or a CD-ROM.

The package insert may also include other information concerning the pharmaceutical compositions and dosage forms in the kit. Generally, such information aids patients and physicians in using the enclosed pharmaceutical compositions and dosage forms effectively and safely. For example, the following information regarding ezetimibe (*e.g.*, Zetia®) and/or simvastatin (*e.g.*, Zocor®) may be supplied in the insert: pharmacokinetics, pharmacodynamics, clinical studies, efficacy parameters, indications and usage, contraindications, warnings, precautions, adverse reactions, overdose, proper dosage and administration, how supplied, proper storage conditions, references and patent information.

The kits of the invention may also include simvastatin (



) preferably combined with a pharmaceutically acceptable carrier, in a pharmaceutical formulation, more preferably in a pharmaceutical dosage form such as a pill, a powder, an injectable liquid, a tablet, dispersible granules, a capsule, a cachet or a suppository. Preferably, the dosage form of simvastatin is a Zocor® tablet (Merck & Co.; Whitehouse Station, NJ).

Tablets or pills comprising simvastatin may be supplied in any convenient form. For example, pills or tablets comprising 5mg simvastatin can be supplied as follows: bottles of 30, 60, 90, 100 or 1000. Pills or tablets comprising 10 mg simvastatin may be supplied as follows: bottles of 30, 60, 90, 100, 1000 or 10,000. Pills or tablets comprising 20 mg simvastatin may be supplied as follows: bottles of 30, 60, 90, 100, 1000 or 10,000. Pills or tablets comprising 40 mg simvastatin may be supplied as follows: bottles of 30, 60, 90, 100 or 1000. Pills or tablets comprising 80 mg simvastatin may be supplied as follows: bottles of 30, 60, 90, 100, 1000 or 10,000.

Ezetimibe and simvastatin may be supplied, in the kit, as separate compositions or combined into a single composition. For example, ezetimibe and simvastatin may be supplied within a single, common pharmaceutical dosage form (e.g., pill or tablet) as in separate pharmaceutical dosage forms (e.g., two separate pills or tablets).

***npc1l1*⁻ Cells**

The present invention provides any isolated mammalian cell, (e.g., an isolated mouse cell, an isolated rat cell or an isolated human cell) which lacks an NPC1L1 gene which encodes or can produce a functional NPC1L1 protein. Included within this embodiment are mutant *npc1l1* genes comprising a point mutation, truncation or deletion of the genetic coding region or of any regulatory element (e.g., a promoter).

For example, the cell can be isolated from a mutant mouse comprising a homozygous mutation of endogenous, chromosomal NPC1L1 wherein the mouse does not produce any functional NPC1L1 protein (e.g., the mouse described below in Example 22). Moreover, the present invention comprises any cell, tissue, organ, fluid, nucleic acid, peptide or other biological substance derived or isolated from such a mutant mouse, particularly a mutant, transgenic mouse which does not produce any functional NPC1L1, wherein the region of endogenous, chromosomal NPC1L1 deleted, in the mouse, corresponds to nucleotides 790-998 of the nucleotide sequence set forth in SEQ ID NO: 45.

The isolated cell can be isolated or derived, for example, from the duodenum, gall bladder, liver, small intestine or stomach of the mutant mouse. Further, the cell can be an enterocyte.

The *npc1l1*⁻ mutant cells are useful, for example, for use in control experiments in screening assays (see e.g., supra) since they lack any NPC1L1-dependent uptake or binding of sterol, 5 α -stanol or ezetimibe. The level of inhibition caused by a particular sample, in a screening assay, can be compared to that of an assay performed with the mutant cell. Ideally, though by no means necessarily, in a screening assay, for example, as described herein, the same amount of binding will be observed by a non-mutant cell or cell membrane, in the presence of an antagonist, as is observed in connection with a mutant *npc1l1* cell or cell membrane alone.

EXAMPLES

The following examples are provided to more clearly describe the present invention and should not be construed to limit the scope of the invention in any way.

5 Example 1: Cloning and Expression of Rat, Mouse and Human NPC1L1.

Rat *NPC1L1*, mouse *NPC1L1* or human *NPC1L1* can all conveniently be amplified using polymerase chain reaction (PCR). In this approach, DNA from a rat, mouse or human cDNA library can be amplified using appropriate primers and standard PCR conditions. Design of primers and optimal amplification conditions
10 constitute standard techniques which are commonly known in the art.

An amplified *NPC1L1* gene may conveniently be expressed, again, using methods which are commonly known in the art. For example, NPC1L1 may be inserted into a pET-based plasmid vector (Stratagene; La Jolla, CA), downstream of the T7 RNA polymerase promoter. The plasmid may then be transformed into a T7
15 expression system (*e.g.*, BL21DE3 *E. coli* cells), grown in a liquid culture and induced (*e.g.*, by adding IPTG to the bacterial culture).

Example 2: Direct Binding Assay.

Membrane preparation: Caco2 cells transfected with an expression vector containing a polynucleotide encoding NPC1L1 (*e.g.*, SEQ ID NO: 2, 4 or 12)
20 are harvested by incubating in 5 mM EDTA/phosphate-buffered saline followed by repeated pipeting. The cells are centrifuged 5 min at 1000 x g. The EDTA/PBS is decanted and an equal volume of ice-cold 50mM Tris-HCl, pH 7.5 is added and cells are broken up with a Polytron (PT10 tip, setting 5, 30 sec). Nuclei and unbroken cells are sedimented at 1000 x g for 10 min and then the supernatant is centrifuged at
25 50,000 x g for 10 min. The supernatant is decanted, the pellet is resuspended by Polytron, a sample is taken for protein assay (bicinchoninic acid, Pierce), and the tissue is again centrifuged at 50,000 x g. Pellets are stored frozen at -20°C.

Binding assay: For saturation binding, four concentrations of [³H]-ezetimibe (15 Ci/mmol) are incubated without and with 10⁻⁵ M ezetimibe in triplicate
30 with 50 µg of membrane protein in a total volume of 200 µl of 50 mM Tris-HCl, pH 7.5, for 30 min at 30°C. Samples are filtered on GF/B filters and washed three times with 2 ml of cold Tris buffer. Filters are dried in a microwave oven, impregnated

with Meltilex wax scintillant, and counted at 45% efficiency. For competition binding assays, five concentrations of a sample are incubated in triplicate with 18 nM [³H]-ezetimibe and 70 µg of membrane protein under the conditions described above. Curves are fit to the data with Prism (GraphPad Software) nonlinear least-squares
5 curve-fitting program and K_i values are derived from IC₅₀ values according to Cheng and Prusoff (Cheng, Y. C., *et al.*, (1973) *Biochem. Pharmacol.* 22: 3099-3108).

Example 3A: SPA Assay.

For each well of a 96 well plate, a reaction mixture of 10 µg human, mouse or rat NPC1L1-CHO overexpressing membranes (Biosignal) and 200 µg/well
10 YSi-WGA-SPA beads (Amersham) in 100 µl is prepared in NPC1L1 assay buffer (25 mM HEPES, pH 7.8, 2 mM CaCl₂, 1mM MgCl₂, 125 mM NaCl, 0.1% BSA). A 0.4 nM stock of ligand- [¹²⁵I]-ezetimibe- is prepared in the NPC1L1 assay buffer. The above solutions are added to a 96-well assay plate as follows: 50 µl NPC1L1 assay buffer, 100 µl of reaction mixture, 50 µl of ligand stock (final ligand concentration is
15 0.1 nM). The assay plates are shaken for 5 minutes on a plate shaker, then incubated for 8 hours before cpm/well are determined in Microbeta Trilux counter (PerkinElmer).

These assays will indicate that [¹²⁵I]-ezetimibe binds to the cell membranes expressing human, mouse or rat NPC1L1. Similar results will be
20 obtained if the same experiment is performed with radiolabeled cholesterol (*e.g.*, ¹²⁵I-cholesterol).

Example 3A: Alternate SPA Assay.

The final concentrations should be: 1 nM ³⁵S-2 (K_m ~ 2-5 nM, ~50,000 dpm/assay);

25 1 µg membranes (~1-2 nM receptor); 0.007%-0.03% taurocholate (0.140 µl 1% stock); 0.010%-0.05% digitonin (0.200 µl 1% stock); 5% DMSO (1.00 µl inhibitors).

In each well of a 96 well plate is put 1 µl DMSO inhibitor solution, and then the radioligand and detergents are added as a 2X solution in 10 µl buffer A.
30 Shake for a minute to be sure the inhibitor and ligand are mixed, before initiating with 9 µl diluted receptor solution in buffer A. After shaking again, the plate is incubated at 37°C for 2 hours. Then WGA beads (0.3 mg) are added as a 3 µl suspension in

buffer A, then shake for 30 minutes. Similar results are obtained if membranes are pre-incubated with beads for 30 minutes before adding ligands. Finally, dilute to 300 μ l with buffer A, cover the plate, spin at 3,000 rpm x 5 min, and read at 2 minutes per well in the "Microbeta" counter.

5 Stocks

Ligand: ^{35}S -2 is 525.42 nM, 0616 $\mu\text{Ci}/\mu\text{l}$, in acetonitrile; spec. act. = 3.8916×10^{-4} fmol/dpm; 1168 Ci/nmol)

Membranes: 3rd batch recombinant human expressed in HEK-293 cells; 20.2 $\mu\text{g}/\mu\text{l}$ stock; ~20–40 pmol NPC1L1/mg protein

10 Buffer A: 26 mM NaHCO_3 ; 0.96 mM NaH_2PO_4 ; 5 mM HEPES; optional addition of 5.5 mM glucose; 117 mM NaCl; 5.4 mM KCl

Example 4: Cholesterol Uptake Assay.

CHO cells expressing either SR-B1 or three different clones of rat NPC1L1 or one clone of mouse NPC1L1 were starved overnight in cholesterol free media then dosed with $[^3\text{H}]$ -cholesterol in a mixed synthetic micelle emulsion for 4 min, 8 min, 12 min or 24 min in the absence or presence of 10 μM ezetimibe. The cells were harvested and the lipids were organically extracted. The extracted lipids were spotted on thin-layer chromatography (TLC) plates and resolved within an organic vapor phase. The free cholesterol bands for each assay were isolated and
20 counted in a scintillation counter.

The SR-B1 expressing cells exhibited an increase in $[^3\text{H}]$ -cholesterol uptake as early as 4 min which was also inhibited by ezetimibe. The three rat clones and the one mouse clone appeared to give background levels of $[^3\text{H}]$ -cholesterol uptake which was similar to that of the untransformed CHO cell.

25 These experiments will yield data demonstrating that CHO cells can perform mouse, rat and human NPC1L1-dependent uptake of $[^3\text{H}]$ -cholesterol when more optimal experimental conditions are developed.

Example 5: Expression of Rat NPC1L1 in Wistar Rat Tissue.

In these experiments, the expression of rat *NPC1L1* mRNA, in several
30 rat tissues, was evaluated. The tissues evaluated were esophagus, stomach, duodenum, jejunum, ileum, proximal colon, distal colon, liver, pancreas, heart, aorta, spleen, lung, kidney, brain, muscle, testes, ovary, uterus, adrenal gland and

thyroid gland. Total RNA samples were isolated from at least 3 male and 3 female animals and pooled. The samples were then subjected to real time quantitative PCR using Taqman analysis using standard dual-labeled fluorogenic oligonucleotide probes. Typical probe design incorporated a 5' reporter dye
 5 (e.g., 6FAM (6-carboxyfluorescein) or VIC) and a 3' quenching dye (e.g., TAMRA (6-carboxytetramethyl-rhodamine)).

rat *NPC1L1*:

Forward: TCTTCACCCTTGCTCTTTGC (SEQ ID NO: 14)
 Reverse: AATGATGGAGAGTAGGTTGAGGAT (SEQ ID NO: 15)
 10 Probe: [6FAM]TGCCCACCTTTGTTGTCTGCTACC[TAMRA]
 (SEQ ID NO: 16) rat β -actin:
 Forward: ATCGCTGACAGGATGCAGAAG (SEQ ID NO: 17)
 Reverse: TCAGGAGGAGCAATGATCTTGA (SEQ ID NO: 18)
 Probe: [VIC]AGATTACTGCCCTGGCTCCTAGCACCAT[TAMRA]
 15 (SEQ ID NO: 19)

PCR reactions were run in 96-well format with 25 μ l reaction mixture in each well containing: Platinum SuperMix (12.5 μ l), ROX Reference Dye (0.5 μ l), 50 mM magnesium chloride (2 μ l), cDNA from RT reaction (0.2 μ l). Multiplex reactions contained gene specific primers at 200 nM each and FAM labeled probe at
 20 100 nM and gene specific primers at 100 nM each and VIC labeled probe at 50 nM. Reactions were run with a standard 2-step cycling program, 95°C for 15 sec and 60°C for 1 min, for 40 cycles.

The highest levels of expression were observed in the duodenum, jejunum and ileum tissue. These data indicate that NPC1L1 plays a role in cholesterol
 25 absorption in the intestine.

Example 6: Expression of Mouse *NPC1L1* in Mouse Tissue.

In these experiments, the expression of mouse *NPC1L1* mRNA, in several tissues, was evaluated. The tissues evaluated were adrenal gland, BM, brain, heart, islets of langerhans, LL, small intestine, kidney, liver, lung, MLN, PLN, muscle,
 30 ovary, pituitary gland, placenta, Peyers Patch, skin, spleen, stomach, testes, thymus, thyroid gland, uterus and trachea. Total RNA samples were isolate from at least 3 male and 3 female animals and pooled. The samples were then subjected to real time quantitative PCR using Taqman analysis using the following primers and probes:

mouse *NPC1L1*:

Forward: ATCCTCATCCTGGGCTTTGC (SEQ ID NO: 20)

Reverse: GCAAGGTGATCAGGAGGTTGA (SEQ ID NO: 21)

Probe: [6FAM]CCCAGCTTATCCA GATTTTCTTCTTCCGC[TAMRA]

5 (SEQ ID NO: 22)

The highest levels of expression were observed in the Peyer's Patch, small intestine, gall bladder and stomach tissue. These data are consistent with a cholesterol absorption role for NPC1L1 which takes place in the digestive system.

Example 7: Expression of Human NPC1L1 in Human Tissue.

10 In these experiments, the expression level of human NPC1L1 mRNA was evaluated in 2045 samples representing 46 normal tissues. Microarray-based gene expression analysis was performed on the Affymetrix HG-U95 GeneChip using a cRNA probe corresponding to base pairs 4192-5117 (SEQ ID NO: 43) in strict accordance to Affymetrix's established protocols. Gene Chips were scanned under
15 low photo multiplier tube (PMT), and data were normalized using either Affymetrix MAS 4.0 or MAS 5.0 algorithms. In addition "spike ins" for most samples were used to construct a standard curve and obtain RNA concentration values according Gene Logic algorithms and procedures. A summary of these results are indicated, below, in Table 2.

Table 2. Expression level of *NPC1L1* mRNA in various human tissues.

Tissue	Present	Absent	Lower 25%	Median	Upper 75%	Tissue	Present	Absent	Lower 25%	Median	Upper 75%
Adipose	2 of 32	30 of 32	-2.45	1.16	12.23	Liver	32 of 34	2 of 34	325.74	427.77	540.1
Adrenal Gland	0 of 12	12 of 12	-23.54	-4.47	10.51	Lung	2 of 93	91 of 93	-3.47	11.03	22.34
Appendix	0 of 3	3 of 3	-9.02	-6.69	38.19	Lymph Node	0 of 11	11 of 11	-1.78	-0.19	1.34
Artery	0 of 3	3 of 3	-6.59	-4.67	9.68	Muscle	0 of 39	39 of 39	-21.57	8.25	26.73
Bladder	1 of 5	4 of 5	-22	-7.95	-1.89	Myometrium	8 of 106	98 of 106	-3.98	4.87	17.55
Bone	0 of 3	3 of 3	-1.84	3.3	19.53	Omentum	0 of 15	15 of 15	-14.25	-1.6	19.58
Breast	4 of 80	76 of 80	-4.07	3.13	14.67	Ovary	1 of 74	73 of 74	0.5	17.51	38.28
Cerebrum	0 of 5	5 of 5	-3.04	3.24	15.38	Pancreas	0 of 34	34 of 34	-87.08	-53.2	-24.14
Cervix	3 of 101	98 of 101	-7.56	-0.07	20.89	Placenta	0 of 5	5 of 5	-20.4	-3.44	18.91
Colon	9 of 151	142 of 151	-10.19	0.91	18.36	Prostate	0 of 32	32 of 32	1.08	13.56	27.24
Cortex Frontal Lobe	0 of 7	7 of 7	1.4	8.46	11.75	Rectum	1 of 43	42 of 43	-9.26	-1.49	9.8
Cortex Temporal Lobe	0 of 3	3 of 3	7.1	8.5	15.87	Right Atrium	4 of 169	165 of 169	-19.32	-8.58	7.72
Duodenum	59 of 61	2 of 61	519.23	927.13	1101.67	Right Ventricle	1 of 160	159 of 160	-24.01	-8.49	10.86
Endometrium	0 of 21	21 of 21	-14.43	-6.39	2.78	Skin	0 of 59	59 of 59	-12.68	1.5	22.77
Esophagus	1 of 27	26 of 27	-10.93	-4.97	12.48	Small Intestine	46 of 63	22 of 66	21.21	493.93	939.2
Fallopian Tube	3 of 51	48 of 51	5.02	13.24	26.77	Soft Tissue	1 of 6	5 of 6	-1.99	2.6	5.32
GallBladder	0 of 0	0 of 0	205.76	273.39	122.3	Spleen	0 of 31	31 of 31	-9.41	-0.31	9.5
Heart	0 of 3	3 of 3	3.33	11.19	11.66	Stomach	7 of 17	40 of 17	19.02	52.29	117.09
Hippocampus	0 of 5	5 of 5	8.25	9.11	19.83	Testis	0 of 5	5 of 5	-4.51	1.22	11.2
Kidney	4 of 86	82 of 86	-8.36	9.41	16.46	Thymus	1 of 71	70 of 71	-6.26	2.51	11.67
Larynx	0 of 4	4 of 4	-13.75	-0.31	8.54	Thyroid Gland	1 of 18	17 of 18	-12.22	2.84	17.86
Left Atrium	2 of 141	139 of 141	-18.9	-4.58	6.84	Uterus	0 of 58	58 of 58	-10.67	1.59	16.01
Left Ventricle	0 of 15	15 of 15	-21.19	-9.59	17.7	WBC	3 of 40	37 of 40	-16.45	-0.72	25.18

Shaded data corresponds to tissues wherein the highest levels of *NPC1L1* mRNA was detected. The "Present" column indicates the proportion of specified tissue samples evaluated wherein *NPC1L1* mRNA was detected. The "Absent" column indicates the proportion of specified tissue samples evaluated wherein *NPC1L1* RNA was not detected. The "lower 25%", "median" and "upper 75%" columns indicate statistical distribution of the relative *NPC1L1* signal intensities observed for each set of tissue evaluated.

Example 8: Distribution of Rat *NPC1L1*, Rat *IBAT* or Rat *SR-B1* mRNA in Rat Small Intestine.

In these experiments, the distribution of rat *NPC1L1* mRNA along the proximal-distal axis of rat small intestines was evaluated. Intestines were isolated from five independent animals and divided into 10 sections of approximately equal length. Total RNA was isolated and analyzed, by real time quantitative PCR using Taqman analysis, for localized expression levels of rat *NPC1L1*, rat *IBAT* (ileal bile acid transporter) or rat *SR-B1* mRNA. The primers and probes used in the analysis were:

rat NPC1L1:

Forward: TCTTCACCCTTGCTCTTTGC (SEQ ID NO: 23)

Reverse: AATGATGGAGAGTAGGTTGAGGAT (SEQ ID NO: 24)

Probe: [6FAM]TGCCACCTTTGTTGTCTGCTACC[TAMRA]

5 (SEQ ID NO: 25) rat Villin:

Forward: AGCACCTGTCCACTGAAGATTTC (SEQ ID NO: 26)

Reverse: TGGACGCTGAGCTTCAGTTCT (SEQ ID NO: 27)

Probe: [VIC]CTTCTCTGCGCTGCCTCGATGGAA[TAMRA] (SEQ

ID NO: 28) rat SR-B1:

10 Forward: AGTAAAAAGGGCTCGCAGGAT (SEQ ID NO: 29)

Reverse: GGCAGCTGGTGACATCAGAGA (SEQ ID NO: 30)

Probe: [6FAM]AGGAGGCCATGCAGGCCTACTCTGA[TAMRA]

(SEQ ID NO: 31) rat IBAT:

Forward: GAGTCCACGGTCAGTCCATGT (SEQ ID NO: 32)

15 Reverse: TTATGAACAACAATGCCAAGCAA (SEQ ID NO: 33)

Probe: [6FAM]AGTCCTTAGGTAGTGGCTTAGTCCCTGGAAGC

TC[TAMRA] (SEQ ID NO: 34)

The mRNA expression levels of each animal intestinal section were analyzed separately, then the observed expression level was normalized to the observed level of villin mRNA in that intestinal section. The observed, normalized mRNA expression levels for each section were then averaged.

The expression level of *NPC1L1* and *SR-B1* were highest in the jejunum (sections 2-5) as compared to that of the more distal ileum sections. Since the jejunum is believed to be the site of cholesterol absorption, these data suggest such a role for rat NPC1L1. *IBAT* distribution favoring the ileum is well documented and served as a control for the experiment.

Example 9: In situ Analysis of Rat NPC1L1 mRNA in Rat Jejunum Tissue.

The localization of rat *NPC1L1* mRNA was characterized by *in situ* hybridization analysis of rat jejunum serial sections. The probes used in this analysis were:

T7-sense probe: GTAATACGACTCACTATAGGGCCCTGACGGT
CCTTCCTGAGGGAATCTTCAC (SEQ ID NO: 35)

T7-antisense probe: GTAATACGACTCACTATAGGGCCTGGGAA
GTTGGTCATGGCCACTCCAGC (SEQ ID NO: 36)

The RNA probes were synthesized using T7 RNA polymerase amplification of a PCR amplified DNA fragment corresponding rat *NPC1L1* nucleotides 3318 to 3672 (SEQ ID NO 1). Sense and anti-sense digoxigenin-UTP labeled cRNA probes were generated from the T7 promoter using the DIG RNA Labeling Kit following the manufacturer's instructions. Serial cryosections rat jejunum were hybridized with the sense and antisense probes. Digoxigenin labeling was detected with the DIG Nucleic Acid Detection Kit based on previous methods. A positive signal is characterized by the deposition of a red reaction product at the site of hybridization.

The anti-sense probe showed strong staining of epithelium along the crypt-villus axis under low magnification (40X). The observed rat *NPC1L1* mRNA expression levels may have been somewhat greater in the crypts than in the villus tips. Under high magnification (200X), staining was observed in the enterocytes but not in the goblet cells. A lack of staining observed with the sense probe (control) confirmed the high specificity of the *NPC1L1* anti-sense signal. These data provided further evidence of the role of rat *NPC1L1* in intestinal cholesterol absorption.

**Example 10: FACS Analysis of Fluorescently Labeled
Ezetimibe Binding to Transiently Transfected CHO Cells.**

In these experiments, the ability of BODIPY-labeled ezetimibe (Altmann, *et al.*, (2002) Biochim. Biophys. Acta 1580(1): 77-93) to bind to *NPC1L1* and *SR-B1* was evaluated. "BODIPY" is a fluorescent group which was used to detect the BODIPY-ezetimibe. Chinese hamster ovary (CHO) cells were transiently transfected with rat *NPC1L1* DNA (rNPC1L1/CHO), mouse *NPC1L1* DNA (mNPC1L1/CHO), mouse *SR-B1* DNA (mSRBI/CHO) or *EGFP* DNA (EGFP/CHO). EGFP is enhanced green fluorescent protein which was used as a positive control. The transfected CHO cells or untransfected CHO cells were then stained with 100 nM BODIPY-labeled ezetimibe and analyzed by FACS. Control experiments were also performed wherein the cells were not labeled with the BODIPY-ezetimibe and wherein untransfected CHO cells were labeled with the BODIPY-ezetimibe.

No staining was observed in the untransfected CHO, rNPC1L1/CHO or mNPC1L1/CHO cells. Fluorescence was detected in the positive-control

EGFP/CHO cells. Staining was also detected in the mouse SR-B1/CHO cells. These data show that, under the conditions tested, BODIPY-ezetimibe is capable of binding to SR-B1 and that such binding is not ablated by the presence of the fluorescent BODIPY group. When more optimal conditions are determined, BODIPY-ezetimibe will be shown to label the rNPC1L1/CHO and mNPC1L1/CHO cells.

Example 11: FACS Analysis of Transiently Transfected CHO Cells Labeled with Anti-FLAG Antibody M2.

In these experiments, the expression of FLAG-tagged NPC1L1 on CHO cells was evaluated. CHO cells were transiently transfected with mouse *NPC1L1* DNA, rat *NPC1L1* DNA, FLAG- rat *NPC1L1* DNA or FLAG- mouse *NPC1L1* DNA. The 8 amino acid FLAG tag used was DYKDDDDK (SEQ ID NO: 37) which was inserted on the amino-terminal extracellular loop just past the secretion signal sequence. The cells were incubated with commercially available anti-FLAG monoclonal mouse antibody M2 followed by a BODIPY-tagged anti-mouse secondary antibody. The treated cells were then analyzed by FACS.

The M2 antibody stained the CHO cells transfected with FLAG-rat *NPC1L1* DNA and with FLAG-mouse NPC1L1. No staining was observed in the CHO cells transfected with mouse *NPC1L1* DNA and with rat *NPC1L1* DNA. These data showed that rat NPC1L1 and mouse NPC1L1 possess no significant, inherent fluorescence and are not bound by the anti-FLAG antibody. The observed, FLAG-dependent labeling of the cells indicated that the FLAG-mouse NPC1L1 and FLAG-rat NPC1L1 proteins are localized at the cell membrane of the CHO cells.

Example 12: FACS Analysis of FLAG-rat NPC1L1-EGFP Chimera in Transiently Transfected CHO Cells.

In these experiments, the surface and cytoplasmic localization of rat NPC1L1 in CHO cells was evaluated. CHO cells were transiently transfected with FLAG- rat *NPC1L1* DNA or with FLAG-rat *NPC1L1-EGFP* DNA. In these fusions, the FLAG tag is at amino-terminus of rat NPC1L1 and EGFP fusion is at the carboxy-terminus of rat NPC1L1. The cells were then stained with the M2 anti-FLAG mouse (primary) antibody followed by secondary staining with a BODIPY-labeled anti-mouse antibody. In control experiments, cells were stained with only the secondary antibody and not with the primary antibody (M2). The stained cells were then analyzed by FACS.

In a control experiment, FLAG-rat NPC1L1 transfected cells were stained with BODIPY anti-mouse secondary antibody but not with the primary antibody. The data demonstrated that the secondary, anti-mouse antibody possessed no significant specificity for FLAG-rat NPC1L1 and that the FLAG-rat NPC1L1, itself, possesses no significant fluorescence.

In another control experiment, unlabeled FLAG-rat NPC1L1-EGFP cells were FACS analyzed. In these experiments, autofluorescence of the enhanced green fluorescent protein (EGFP) was detected.

FLAG-rat NPC1L1 cells were stained with anti-FLAG mouse antibody M2 and with the BODIPY-labeled anti-mouse secondary antibody and FACS analyzed. The data from this analysis showed that the cells were labeled with the secondary, BODIPY-labeled antibody which indicated expression of the FLAG-rat NPC1L1 protein on the surface of the CHO cells.

FLAG-rat NPC1L1-EGFP cells were stained with anti-FLAG mouse antibody M2 and with the BODIPY-labeled anti-mouse secondary antibody and FACS analyzed. The data from this analysis showed that both markers (BODIPY and EGFP) were present indicating surface expression of the chimeric protein. The data also indicated that a portion of the protein was located within the cells and may be associated with transport vesicles. These data supported a role for rat NPC1L1 in vesicular transport of cholesterol or protein expressed in subcellular organelles such as the rough endoplasmic reticulum.

Example 13: FACS Analysis and Fluorescent Microscopy of FLAG-rat NPC1L1-EGFP Chimera in a Cloned CHO Cell Line.

In these experiments, the cellular localization of rat NPC1L1 was evaluated by FACS analysis and by immunohistochemistry. CHO cells were transfected with FLAG-rat *NPC1L1-EGFP* DNA and stained with anti-FLAG mouse antibody M2 and then with a BODIPY-labeled anti-mouse secondary antibody. In the fusion, the FLAG tag is at the amino-terminus of rat NPC1L1 and the enhanced green fluorescent protein (EGFP) tag is located at the carboxy-terminus of the rat NPC1L1. The stained cells were then analyzed by FACS and by fluorescence microscopy.

Cells transfected with FLAG-rat *NPC1L1-EGFP* DNA were stained with the anti-FLAG mouse antibody M2 and then with the BODIPY-labeled anti-

mouse secondary antibody. FACS analysis of the cells detected both markers indicating surface expression of the chimeric protein.

FLAG-rat NPC1L1-EGFP transfected cells were analyzed by fluorescent microscopy at 63X magnification. Fluorescent microscopic analysis of the cells indicated non-nuclear staining with significant perinuclear organelle staining. Resolution of the image could not confirm the presence of vesicular associated protein. These data indicated that the fusion protein was expressed on the cell membrane of CHO cells.

Example 14: Generation of Polyclonal Anti-rat NPC1L1 Rabbit Antibodies.

Synthetic peptides (SEQ ID NO: 39-42) containing an amino- or carboxy-terminal cysteine residue were coupled to keyhole limpet hemocyanin (KLH) carrier protein through a disulfide linkage and used as antigen to raise polyclonal antiserum in New Zealand white rabbits (range 3-9 months in age). The KLH-peptide was emulsified by mixing with an equal volume of Freund's Adjuvant, and injected into three subcutaneous dorsal sites. Prior to the 16 week immunization schedule a pre-immune sera sample was collected which was followed by a primary injection of 0.25 mg KLH-peptide and 3 scheduled booster injections of 0.1 mg KLH-peptide. Animals were bled from the auricular artery and the blood was allowed to clot and the serum was then collected by centrifugation

The anti-peptide antibody titer was determined with an enzyme linked immunosorbent assay (ELISA) with free peptide bound in solid phase (1µg/well). Results are expressed as the reciprocal of the serum dilution that resulted in an OD₄₅₀ of 0.2. Detection was obtained using the biotinylated anti-rabbit IgG, horse radish peroxidase-streptavidin (HRP-SA) conjugate, and ABTS.

Example 15: FACS Analysis of Rat NPC1L1 Expression in CHO Cells Transiently Transfected with Rat NPC1L1 DNA Using Rabbit Anti-rat NPC1L1 Antisera.

In these experiments, the expression of rat NPC1L1 on the surface of CHO cells was evaluated. CHO cells were transfected with rat *NPC1L1* DNA, then incubated with either rabbit preimmune serum or with 10 week anti-rat NPC1L1 serum described, above, in Example 14 (i.e., A0715, A0716, A0867 or A0868). Cells labeled with primary antisera were then stained with a BODIPY-modified anti-rabbit secondary antibody followed by FACS analysis.

No antibody surface labeling was observed for any of the pre-immune sera samples. Specific cell surface labeling of rat NPC1L1 transfected cells was observed for both A0715 and A0868. Antisera A0716 and A0867 did not recognize rat NPC1L1 surface expression in this assay format. This indicates that the native,
5 unfused rat NPC1L1 protein is expressed in the CHO cells and localized to the CHO cell membranes. Cell surface expression of NPC1L1 is consistent with a role in intestinal cholesterol absorption.

Example 16: FACS Analysis of CHO Cells Transiently Transfected with FLAG-Mouse NPC1L1 DNA or FLAG-rat NPC1L1 DNA or Untransfected CHO Cells Using Rabbit Anti-rat NPC1L1 Antisera.
10

In these experiments, the expression of FLAG-mouse NPC1L1 and FLAG-rat NPC1L1 in CHO cells was evaluated. CHO cells were transiently transfected with FLAG-mouse *NPC1L1* DNA or with FLAG-rat *NPC1L1* DNA. The FLAG-mouse NPC1L1 and FLAG-rat NPC1L1 transfected cells were labeled with
15 either A0801, A0802, A0715 or A0868 sera (see Example 14) or with anti-FLAG antibody, M2. The labeled cells were then stained with BODIPY-labeled anti-rabbit secondary antibody and FACS analyzed. The untransfected CHO cells were analyzed in the same manner as the transfected cell lines.

Positive staining of the untransfected CHO cells was not observed for
20 any of the antisera tested. Serum A0801-dependent labeling of FLAG-rat NPC1L1 transfected cells was observed but such labeling of FLAG-mouse NPC1L1 transfected cells was not observed. Serum A0802-dependent labeling of FLAG-mouse NPC1L1 or FLAG-rat NPC1L1 transfected cells was not observed. Strong serum A0715-dependent labeling of FLAG-rat NPC1L1 transfected cells was observed and weak
25 serum A0715-dependent labeling of FLAG-mouse NPC1L1 transfected cells was observed. Weak serum A0868-dependent labeling of rat NPC1L1 and mouse NPC1L1 transfected cells was observed. Strong Anti-FLAG M2 antibody-dependent labeling of FLAG-rat NPC1L1 and FLAG-mouse NPC1L1 transfected cells was observed. The strong M2 staining is likely to be due to the fact that M2 is an affinity-
30 purified, monoclonal antibody of known concentration. In contrast, the respective antisera are polyclonal, unpurified and contain an uncertain concentration of anti-rat NPC1L1 antibody. These data provide further evidence that the FLAG-mouse NPC1L1 and FLAG-rat NPC1L1 proteins are expressed in CHO cells and localized to

the CHO cell membranes. Cell surface expression of NPC1L1 is consistent with a role in intestinal cholesterol absorption.

Example 17: Immunohistochemical Analysis of Rat Jejunum Tissue with Rabbit Anti-rat NPC1L1 Antisera A0715.

5 In these experiments, the localization of rat NPC1L1 in rat jejunum was analyzed by immunohistochemistry. Rat jejunum was removed, immediately embedded in O.C.T. compound and frozen in liquid nitrogen. Sections (6µm) were cut with a cryostat microtome and mounted on glass slides. Sections were air dried at room temperature and then fixed in Bouin's fixative. Streptavidin-biotin-peroxidase
10 immunostaining was carried out using Histostain-SP kit. Endogenous tissue peroxidase activity was blocked with a 10 minute incubation in 3% H₂O₂ in methanol, and nonspecific antibody binding was minimized by a 45 minute incubation in 10% nonimmune rabbit serum. Sections were incubated with a rabbit anti-rat NPC1L1 antisera A0715 or A0868 at a 1: 500 dilution at 4°C, followed by incubation with
15 biotinylated goat anti-rabbit IgG and with streptavidin-peroxidase. Subsequently, the sections were developed in an aminoethyl carbazole (AEC)-H₂O₂ staining system and counterstained with hematoxylin and examined by microscopy. A positive reaction using this protocol is characterized by the deposition of a red reaction product at the site of the antigen-antibody reaction. Nuclei appeared blue from the hematoxylin
20 counterstain. Controls were performed simultaneously on the neighboring sections from the same tissue block. Control procedures consisted of the following: (1) substitute the primary antibody with the pre-immune serum, (2) substitute the primary antibody with the non-immune rabbit serum, (3) substitute the primary antibody with PBS, (4) substitute the second antibody with PBS.

25 The example shows tissue stained with anti-rat NPC1L1 sera A0715 or with the preimmune sera analyzed at low magnification (40X) and at high magnification (200X). The A0715-stained tissue, at low magnification, showed positive, strong staining of the villi epithelial layer (enterocytes). The A0715-stained tissue at high magnification showed positive, strong staining of the enterocyte apical
30 membranes. No staining was observed in tissue treated only with preimmune sera. Similar results were obtained with sera A0868. These data indicate that rat NPC1L1 is expressed in rat jejunum which is consistent with a role in intestinal cholesterol absorption.

Example 18: Labeled Cholesterol Uptake Assay.

In this example, the ability of CHO cells stably transfected with rat NPC1L1 to take up labeled cholesterol was evaluated. In these assays, cholesterol uptake, at a single concentration, was evaluated in a pulse-chase experiment. The data generated in these experiments are set forth, below, in Table 3.

Cells:

- A. CHO cells stably transfected with rat *NPC1L1* cDNA
- B. CHO background (no transfection)

Cells were seeded at 500,000 cells/ well (mL) in 12-well plates.

10 Procedure:

All reagents and culture plates were maintained at 37°C unless otherwise noted.

Starve. The maintenance media (F12 HAMS, 1%Pen/Strep, 10%FCS) was removed and the cells were rinsed with serum-free HAMS media. The serum-free media was then replaced with 1 mL "starve" media (F12 HAMS, Pen/Strep, 5% lipoprotein deficient serum (LPDS)).

One plate of each cell line was starved overnight. The remaining 2 plates were designated "No Starve" (see below).

Pre-Incubation. Media was removed from all plates, rinsed with serum-free HAMS and replaced with starve media for 30 minutes.

³H-Cholesterol Pulse. The following was added directly to each well.

0.5μCi ³H-cholesterol (~1.1 X 10⁶ dpm/well) in 50μl of a mixed bile salt micelle.

4.8 mM sodium taurocholate (2.581 mg/mL)

0.6 mM sodium oleate (0.183 mg/mL)

0.25 mM cholesterol (0.1 mg/mL)

Dispersed in "starve" media by ultrasonic vibration

Final media cholesterol concentration = 5μg/mL

Labeled cholesterol pulse time points were 0, 4, 12 and 24 minutes.

30 Triplicate wells for each treatment were prepared.

Wash. At the designated times, media was aspirated and the cells were washed once with Hobbs Buffer A (50mM Tris, 0.9% NaCl, 0.2% BSA, pH 7.4) and once with Hobbs Buffer B (50mM Tris, 0.9% NaCl, pH 7.4 (no BSA)) at 37°C.

Processing/Analysis. Cells were digested overnight with 0.2N NaOH, 2mL/well at room temperature. One 1.5 mL aliquot was removed from each well, neutralized & counted for radioactivity by scintillation counting. Two additional 50µl aliquots from all wells are assayed for total protein by the Pierce micro BCA method.

5 The quantity of labeled cholesterol observed in the cells was normalized by the quantity of protein in the cells.

Table 3. Uptake of ³H-cholesterol by CHO cells transfected with rat *NPC1L1* or mouse *SR-B1* or untransfected CHO cells.

Time, min		Total Cholesterol, dpm/protein ± sem				Total Cholesterol, dpm/mg protein ± sem			
		NPC1L1		CHO		NPC1L1		CHO	
After ³ H-Cholesterol	0	2067	±46	4568	±1937	10754	±166	22881	±9230
	4	2619	±130	2868	±193	15366	±938	15636	±1471
	12	2868	±193	4459	±170	15636	±1471	24622	±966
	24	7010	±89	7204	±173	41129	±685	39361	±1207
No Starve									
After ³ H-Cholesterol	0	1937	±273	2440	±299	10909	±1847	12429	±1673
	4	3023	±308	2759	±105	17278	±1650	14307	±781
	12	2759	±105	4857	±186	14307	±781	26270	±1473
	24	6966	±72	7344	±65	39196	±174	38381	±161
Starve									

dpm=disintegrations per minute
sem=standard error of the mean

10

Example 19: Effect of Ezetimibe on Cholesterol Uptake.

The effect of ezetimibe on the ability of CHO cells stably transfected with mouse or rat *NPC1L1* or mouse *SR-B1* to take up ³H-labeled cholesterol was evaluated in pulse-chase experiments. One cDNA clone of mouse *NPC1L1* (C7) and three clones of rat *NPC1L1* (C7, C17 and C21) were evaluated. The ability of CHO

15

cells stably transfected with mouse *SR-BI*, mouse *NPC1L1* and rat *NPC1L1* to take up labeled cholesterol, in the absence of ezetimibe, was also evaluated in the pulse-chase experiments. Data generated in these experiments are set forth, below, in Tables 4 and 5. Additionally, the quantity of total cholesterol taken up by transfected and
 5 untransfected CHO cells in the presence of four different unlabeled cholesterol concentrations was also evaluated. The data from these experiments is set forth, below, in Table 6.

Cells:

- A. CHO cells stably transfected with rat or mouse *NPC1L1* cDNA
 - 10 B. CHO background (no transfection)
 - C. *SR-BI* transfected CHO cells
- Cells seeded at 500,000 cells / well (mL) in 12-well plates.

Procedure:

All reagents and culture plates were maintained at 37°C unless
 15 otherwise noted.

Starve. The maintenance media (F12 HAMS, 1%Pen/Strep, 10%FCS) was removed and the cells were rinsed with serum-free HAMS media. The serum-free media was then replaced with 1 mL "starve" media (F12 HAMS, Pen/Strep, 5% lipoprotein deficient serum (LPDS). The cells were then starved overnight.

20 **Pre-Incubation/ pre-dose.** Media was removed from all plates and replaced with fresh starve media and preincubated for 30 minutes. Half of the wells received media containing ezetimibe (stock soln in EtOH; final conc. = 10µM).

³H-Cholesterol Pulse. The following was added directly to each well:

25 0.5µCi ³H-cholesterol (~1.1 X 10⁶ dpm/well) in 50µl of a mixed bile salt micelle

4.8mM sodium taurocholate (2.581mg/mL)

0.6 mM sodium oleate (0.183 mg/mL)

0.25 mM cholesterol (0.1 mg/mL)

Dispersed in "starve" media by ultrasonic vibration

30 Final media cholesterol concentration = 5µg/mL

Labeled cholesterol pulse time points were 4, 12, 24 minutes and 4 hours. Triplicate wells were prepared for each treatment.

Wash. At designated times, media was aspirated and cells were washed once with Hobbs Buffer A (50mM Tris, 0.9% NaCl, 0.2% bovine serum albumin (BSA), pH 7.4) and once with Hobbs Buffer B (50mM Tris, 0.9% NaCl, pH 7.4 (no BSA)) at 37°C.

5 **Processing/Analysis.**

A. 4, 12, 24 minute time points: Cells were digested overnight with 0.2N NaOH, 2mL/well, room temperature. One 1.5 mL aliquot was removed from each well, neutralized & counted for radioactivity by scintillation counting.

10 B. 4 hour time point: The digested cells were analyzed by thin-layer chromatography to determine the content of cholesterol ester in the cells.

 Extracts were spotted onto TLC plates and run for 30 minutes in 2 ml hexane: isopropanol (3: 2) mobile phase for 30 minutes, followed by a second run in 1 ml hexane: isopropanol (3: 2) mobile phase for 15 minutes.

15 C. Protein determination of cell extracts. Plates containing a sample of the cell extracts were placed on orbital shaker at 120 rpm for indicated times and then extracts are pooled into 12 X 75 tubes. Plates were dried and NaOH (2ml/well) added. The protein content of the samples were then determined. Two additional 50µl aliquots from all wells were assayed for total protein by the Pierce micro BCA method. The quantity of labeled cholesterol observed in the cells was
20 normalized to the quantity of protein in the cells.

Table 4. Total Cholesterol in Transfected CHO Cells in the Presence and Absence of Ezetimibe.

Clones:	Total Cholesterol, dpm \pm sem				Total Cholesterol, dpm/mg protein \pm sem			
	Vehicle		EZ (10 μ M)		Vehicle		EZ (10 μ M)	
	4 Min Pulse							
CHO Control	3413	\pm 417	3222	\pm 26	33443	\pm 4070	31881	\pm 483
SR-BI	14207	\pm 51	10968	\pm 821	118242	\pm 1261	92474	\pm 2902
mNPCIL1(C7)	4043	\pm 419	4569	\pm 222	30169	\pm 3242	30916	\pm 1137
rNPCIL1(C21)	3283	\pm 288	3769	\pm 147	23728	\pm 2111	27098	\pm 689
rNPCIL1(C17)	3188	\pm 232	3676	\pm 134	24000	\pm 832	28675	\pm 527
rNPCIL1(C7)	1825	\pm 806	3268	\pm 121	15069	\pm 6794	27285	\pm 968
	12 Min Pulse							
CHO Control	4710	\pm 246	4532	\pm 165	44208	\pm 2702	43391	\pm 1197
SR-BI	16970	\pm 763	12349	\pm 298	140105	\pm 6523	98956	\pm 4447
mNPCIL1(C7)	6316	\pm 85	6120	\pm 755	45133	\pm 342	41712	\pm 4054
rNPCIL1(C21)	5340	\pm 12	4703	\pm 231	40018	\pm 1181	33985	\pm 1928
rNPCIL1(C17)	4831	\pm 431	4579	\pm 257	37378	\pm 3461	34063	\pm 1619
rNPCIL1(C7)	4726	\pm 272	4664	\pm 63	39100	\pm 2350	38581	\pm 784
	24 Min Pulse							
CHO Control	7367	\pm 232	6678	\pm 215	65843	\pm 1281	61764	\pm 2131
SR-BI	39166	\pm 2152	23558	\pm 1310	324126	\pm 11848	198725	\pm 11713
mNPCIL1(C7)	10616	\pm 121	9749	\pm 482	77222	\pm 1040	74041	\pm 3670
rNPCIL1(C21)	9940	\pm 587	8760	\pm 293	76356	\pm 9618	66165	\pm 2181
rNPCIL1(C17)	8728	\pm 721	8192	\pm 237	70509	\pm 5189	62279	\pm 4352
rNPCIL1(C7)	8537	\pm 148	7829	\pm 204	72134	\pm 1305	63482	\pm 368

EZ = ezetimibe

Table 5. Cholesterol Ester in CHO cells in the Presence or Absence of Ezetimibe.

Clones:	Cholesteryl Ester, dpm \pm sem				Cholesteryl Ester, dpm/mg protein \pm sem			
	Vehicle		EZ (10 μ M)		Vehicle		EZ (10 μ M)	
	4 Hour Pulse							
CHO Control	652	± 13	208	± 9	5647	± 55	1902	± 87
SR-BI	47608	± 1292	9305	± 401	391067	± 14391	72782	± 3181
mNPCIL1(C7)	732	± 127	453	± 118	4994	± 827	3057	± 776
rNPCIL1(C21)	2667	± 90	454	± 33	18655	± 1032	3193	± 265
rNPCIL1(C17)	751	± 74	202	± 10	5379	± 481	1510	± 62
rNPCIL1(C7)	462	± 25	191	± 54	3597	± 193	1496	± 403

	Free Cholesterol, dpm \pm sem				Free Cholesterol, dpm/mg protein \pm sem			
	Vehicle		EZ (10 μ M)		Vehicle		EZ (10 μ M)	
	4 Hour Pulse							
CHO Control	61612	± 1227	56792	± 568	533876	± 17770	519607	± 16203
SR-BI	214678	± 4241	194519	± 474	1762873	± 46607	1521341	± 4185
mNPCIL1(C7)	79628	± 793	77516	± 1910	544661	± 1269	523803	± 10386
rNPCIL1(C21)	71352	± 1343	69106	± 711	498016	± 8171	485460	± 4410
rNPCIL1(C17)	78956	± 3782	71646	± 446	566456	± 29204	536651	± 7146
rNPCIL1(C7)	75348	± 2093	70628	± 212	586127	± 13932	556855	± 7481

EZ =ezetimibe

Table 6. Uptake of labeled cholesterol in the presence of increasing amounts of unlabeled cholesterol.

	Total Cholesterol, dpm \pm sem				Total Cholesterol, dpm/mg protein \pm sem			
	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)
24 Min Pulse								
Cold Cholesterol								
3 μ g/mL	12271 \pm 430	49603 \pm 2428	14250 \pm 1628	10656 \pm 1213	108936 \pm 4313	541562 \pm 13783	140764 \pm 14433	94945 \pm 12916
10 μ g/mL	16282 \pm 2438	79967 \pm 8151	25465 \pm 3037	13225 \pm 4336	151283 \pm 23345	880224 \pm 82254	250985 \pm 27411	123433 \pm 34092
30 μ g/mL	14758 \pm 1607	71925 \pm 8863	19001 \pm 1530	13218 \pm 1149	135109 \pm 12106	796236 \pm 18952	180435 \pm 12112	111522 \pm 6941
100 μ g/mL	16458 \pm 1614	58185 \pm 4348	15973 \pm 1663	11560 \pm 1132	149559 \pm 17977	630143 \pm 3718	147717 \pm 8261	101328 \pm 7191
	Cholesteryl Ester, dpm \pm sem				Cholesteryl Ester, dpm/mg protein \pm sem			
	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)
4 Hour Pulse								
3 μ g/mL	2737 \pm 114	39596 \pm 1241	1561 \pm 1	4015 \pm 47	22050 \pm 978	382641 \pm 5935	13684 \pm 217	32020 \pm 641
10 μ g/mL	1646 \pm 76	17292 \pm 362	998 \pm 36	1866 \pm 33	13323 \pm 606	157914 \pm 3400	8917 \pm 467	14849 \pm 127
30 μ g/mL	970 \pm 46	6642 \pm 153	537 \pm 82	970 \pm 9	7627 \pm 325	63547 \pm 1760	4885 \pm 748	7741 \pm 100
100 μ g/mL	895 \pm 156	4777 \pm 27	405 \pm 7	777 \pm 16	7133 \pm 1230	45088 \pm 1526	3563 \pm 68	6805 \pm 198
	Free Cholesterol, dpm \pm sem				Free Cholesterol, dpm/mg protein \pm sem			
	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)
4 Hour Pulse								
3 μ g/mL	89013 \pm 3724	211783 \pm 3268	104343 \pm 2112	92244 \pm 987	717308 \pm 34130	2047695 \pm 16213	914107 \pm 5869	735498 \pm 11209
10 μ g/mL	136396 \pm 8366	278216 \pm 10901	196173 \pm 4721	125144 \pm 877	1105118 \pm 76074	2548130 \pm 92471	1753072 \pm 86578	996824 \pm 27850
30 μ g/mL	131745 \pm 2922	224429 \pm 2536	149172 \pm 19089	117143 \pm 4976	1036195 \pm 21142	2149315 \pm 78068	1357136 \pm 180264	934772 \pm 43202
100 μ g/mL	79336 \pm 4011	231470 \pm 4221	114599 \pm 2803	93538 \pm 1588	632965 \pm 29756	2182022 \pm 36793	1035979 \pm 30329	723225 \pm 21694
	Cholesteryl Ester, dpm \pm sem				Cholesteryl Ester, dpm/mg protein \pm sem			
	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)
24 Hour Pulse								
3 μ g/mL	57373 \pm 2704	162296 \pm 1644	22986 \pm 940	59377 \pm 953	357629 \pm 14639	1248900 \pm 18563	160328 \pm 6363	401315 \pm 3557
10 μ g/mL	33730 \pm 1296	112815 \pm 373	14836 \pm 552	31797 \pm 525	215004 \pm 5942	830231 \pm 12764	98594 \pm 4203	200451 \pm 5239
30 μ g/mL	19193 \pm 100	58608 \pm 1413	8878 \pm 355	18963 \pm 380	122071 \pm 1271	446381 \pm 3472	59091 \pm 3697	119728 \pm 2131
100 μ g/mL	16761 \pm 198	51280 \pm 1270	8784 \pm 946	14933 \pm 111	103235 \pm 1739	272796 \pm 13392	60670 \pm 4597	96215 \pm 1023
	Free Cholesterol, dpm \pm sem				Free Cholesterol, dpm/mg protein \pm sem			
	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)	CHO Control	SR-BI	mNPC1L1(C7)	rNPC1L1(C21)
24 Hour Pulse								
3 μ g/mL	248985 \pm 4207	357819 \pm 4519	285610 \pm 5187	227244 \pm 1016	1552637 \pm 18954	2752957 \pm 24984	1993256 \pm 36968	1536023 \pm 10304
10 μ g/mL	231208 \pm 8927	269822 \pm 5872	311777 \pm 8227	231666 \pm 6158	1477414 \pm 25954	1984473 \pm 18420	2068980 \pm 25517	1461157 \pm 38517
30 μ g/mL	203566 \pm 6008	225273 \pm 5932	279604 \pm 6612	209372 \pm 3386	1294878 \pm 41819	1716066 \pm 52381	1839476 \pm 29507	1321730 \pm 5452
100 μ g/mL	178424 \pm 2379	167082 \pm 2211	229832 \pm 4199	182678 \pm 7709	1099648 \pm 25150	1455799 \pm 9385	1599244 \pm 76938	1177546 \pm 51191

Example 20: Labeled Cholesterol Uptake Assay.

5

In this example, the ability of CHO cells transiently transfected with rat *NPC1L1* or mouse *SR-BI* to take up labeled cholesterol was evaluated. Also evaluated was the ability of rat *NPC1L1* to potentiate the ability of CHO cells transfected with mouse *SR-BI* to take up labeled cholesterol. In these assays,

cholesterol uptake, at a single concentration, was evaluated in pulse-chase experiments. The data generated in these experiments are set forth, below, in Table 7.

Cells:

- A. CHO background cells (mock transfection).
- 5 B. CHO cells transiently transfected with mouse SR-B1.
- C. CHO transiently transfected with rat *NPC1L1* cDNAs (n=8 clones).

Transiently transfected cells were seeded at 300,000 cells / well (mL) in 12-well plates.

10 **Procedure:**

All reagents and culture plates were maintained at 37°C unless otherwise noted.

- Starve. The maintenance media (F12 HAMS, 1%Pen/Strep, 10%FCS) was removed from the cells and replaced with 1 mL "starve" media (F12 HAMS, 15 Pen/Strep, 5% lipoprotein deficient serum (LPDS). Cells were starved for 1 hour.

³H-Cholesterol Pulse. The following was added directly to each well.

- 0.5μCi ³H-cholesterol (~1.1 X 10⁶ dpm/well) in 50μl of a mixed bile salt micelle.
- 4.8mM sodium taurocholate (2.581mg/mL)
- 20 0.6 mM sodium oleate (0.183 mg/mL)
- 0.25 mM cholesterol (0.1 mg/mL)
- Dispersed in "starve" media by ultrasonic vibration
- Final media cholesterol concentration = 5μg/mL

Labeled cholesterol pulse time points were 24 Min and 4 hours.

- 25 Triplicate wells for each treatment.

Wash. At the designated times, media was aspirated and cells were washed once with Hobbs Buffer A (50mM Tris, 0.9% NaCl, 0.2% BSA, pH 7.4) and once with Hobbs Buffer B (50mM Tris, 0.9% NaCl, pH 7.4 (no BSA)) at 37°C.

Processing/Analysis.

- 30 A. 24 minute time point: Cells were digested overnight with 0.2N NaOH, 2mL/well at room temp. One 1.5 mL aliquot was removed from each well, neutralized & counted for radioactivity by scintillation counting.

B. 4 hour time point: The digested cells were analyzed by thin-layer chromatography to determine the content of cholesterol ester in the cells.

The extracts were spotted onto thin layer chromatography plates and run in 2 ml hexane: isopropanol (3: 2) containing mobile phase for 30 minutes,
5 followed by a second run in 1 ml hexane: isopropanol (3: 2) containing mobile phase for 15 min.

C. Protein determination of cell extracts: Plates containing a sample of the cell extracts were placed on orbital shaker at 120 rpm for indicated times and then extracts are pooled into 12X75 tubes. Plates were dried and NaOH
10 (2ml/well) added. The protein content of the samples were then determined. Two additional 50µl aliquots from all wells were assayed for total protein by the Pierce micro BCA method. The quantity of labeled cholesterol observed in the cells was normalized to the quantity of protein in the cells.

Table 7. Labeled cholesterol uptake in transiently transfected CHO cells.

Transfection	Total Cholesterol, \pm sem	
	dpm	dpm/mg protein
	24 Min Pulse	
CHO Control (mock)	4721 \pm 436	49024 \pm 4328
SR-BI(Transient)	5842 \pm 82	59445 \pm 1099
NPC1L1 (Transient)	4092 \pm 377	47026 \pm 2658
SR-BI/NPC1L1 (trans)	3833 \pm 158	52132 \pm 3071
	Cholesteryl Ester, \pm sem	
	dpm	dpm/mg protein
	4 Hour Pulse	
CHO Control (mock)	2132 \pm 40	20497 \pm 640
SR-BI(Transient)	5918 \pm 237	51812 \pm 1417
NPC1L1 (Transient)	1944 \pm 93	19788 \pm 642
SR-BI/NPC1L1 (trans)	4747 \pm 39	58603 \pm 1156
	Free Cholesterol, \pm sem	
	dpm	dpm/mg protein
	4 Hour Pulse	
CHO Control (mock)	45729 \pm 328	439346 \pm 5389
SR-BI(Transient)	50820 \pm 2369	444551 \pm 9785
NPC1L1 (Transient)	39913 \pm 1211	406615 \pm 6820
SR-BI/NPC1L1 (trans)	37269 \pm 1225	459509 \pm 6195

Example 21: Expression of rat, mouse and human NPC1L1.

5 In this example, *NPC1L1* was introduced into cells and expressed. Species specific NPC1L1 expression constructs were cloned into the plasmid pCDNA3 using clone specific PCR primers to generate the ORF flanked by appropriate restriction sites compatible with the polylinker of the vector. For all three species of NPC1L1, small intestine total tissue RNA was used as a template for

10 reverse transcriptase-polymerase chain reaction (RT-PCR) using oligo dT as the

template primer. The rat *NPC1L1* was cloned as an EcoRI fragment, human *NPC1L1* was cloned as a XbaI/NotI fragment and mouse *NPC1L1* was cloned as an EcoRI fragment. Forward and reverse strand sequencing of each clone was performed to confirm sequence integrity. Standard transient transfection procedures were used with CHO cells. In a 6-well plate CHO cells were plated 1 day before transfection at a plating density of 2×10^5 cells/well. The following day, cells were incubated with 2 μ g plasmid DNA and 6 μ L Lipofectamine for 5 hours followed a fresh media change. Forty-eight hours later, cells were analyzed for NPC1L1 expression using anti-NPC1L1 antisera by either FACS or western blot. To establish stable long term cell lines expressing NPC1L1, transfected CHO cells were selected in the presence of geneticin (G418, 0.8 mg/ml) as recommended by the manufacturer (Life Technologies). Following one month of selection in culture, the cell population was stained with anti-NPC1L1 antisera and sorted by FACS. Individual positive staining cells were cloned after isolation by limiting dilution and then maintained in selective media containing geneticin (0.5 mg/ml).

Other cell types less susceptible to transfection procedures have been generated using adenoviral vector systems. This system used to express NPC1L1 is derived from Ad 5, a type C adenovirus. This recombinant replication-defective adenoviral vector is made defective through modifications of the E1, E2 and E4 regions. The vector also has additional modifications to the E3 region generally affecting the E3b region genes RIDa and RIDb. NPC1L1 expression was driven using the CMV promoter as an expression cassette substituted in the E3 region of the adenovirus. Rat and mouse NPC1L1 were amplified using clone specific primers flanked by restriction sites compatible with the adenovirus vector. Adenovirus infective particles were produced from 293-D22 cells in titers of 5×10^{10} P/mL. Viral lysates were used to infect cells resistant to standard transfection methodologies. In Caco2 cells, which are highly resistant to heterologous protein expression, adenovirus mediated expression of NPC1L1 has been shown by western blot analysis to persist at least 21 days post-infection.

Example 22: NPC1L1 Knock-Out Transgenic Mouse.

NPC1L1 knockout mice were constructed via targeted mutagenesis. This methodology utilized a targeting construct designed to delete a specific region of the mouse *NPC1L1* gene. During the targeting process the *E. coli* lacZ reporter gene

was inserted under the control of the endogenous *NPC1L1* promoter. The region in *NPC1L1* (SEQ ID NO: 45) being deleted is from nucleotide 790 to nucleotide 998. The targeting vector contains the LacZ-Neo cassette flanked by 1.9 kb 5' arm ending with nucleotide 789 and a 3.2 kb 3' arm starting with nucleotide 999. Genomic DNA

5 from the recombinant embryonic stem cell line was assayed for homologous recombination using PCR. Amplified DNA fragments were visualized by agarose gel electrophoresis. The test PCRs employed a gene specific primer, which lies outside of and adjacent to the targeting vector arm, paired with one of three primers specific to the LacZ-Neo cassette sequence. For 5' PCR reconfirmation, the *NPC1L1* specific

10 oligonucleotide ATGTTAGGTGAGTCTGAACCTACCC (SEQ ID NO: 46) and for 3' PCR reconfirmation the *NPC1L1* specific oligonucleotide GGATTGCATTTCTTCAAGAAAGCC (SEQ ID NO: 47) were used. Genotyping of the F2 mice was performed by multiplex PCR using the *NPC1L1* specific forward primer TATGGCTCTGCCCTCTGCAATGCTC (SEQ ID NO: 48) the LacZ-Neo

15 cassette specific forward primer TCAGCAGCCTCTGTTCCACATACTTC (SEQ ID NO: 49) in combination with the *NPC1L1* gene specific reverse primer GTTCCACAGGGTCTGTGGTGAGTTC (SEQ ID NO: 50) allowed for determination of both the targeted and endogenous alleles. Analysis of the PCR products by agarose gel electrophoresis distinguished the wild-type, heterozygote and

20 homozygote null mouse from each other.

Example 23: Acute Cholesterol Absorption in *NPC1L1*-Deficient Mice.

To determine whether *NPC1L1* plays a role in cholesterol absorption, *NPC1L1* deficient mice were studied.

Mice deficient in *NPC1L1* (-/-) were generated by breeding

25 heterozygote mice (+/-) to obtain wild-type (+/+) and *NPC1L1* deficient mice (-/-). Non-fasted mice (6.5-9 weeks old, mixed 129 and C57BL/6 background) were weighed and grouped (n=2 -/- and n=4 +/+). All animals were gavaged (Feeding needles, 24G x 1 inch, Popper and Sons, NY) with 0.1 ml corn oil (Sigma; St. Louis, MO) containing 1 μ Ci 14 C-cholesterol (New England Nuclear, [14 C] Cholesterol,

30 NEC-018) and 0.1mg carrier cholesterol mass (Sigma; St. Louis, MO). Two hours later, blood was collected by heart puncture. The liver was removed, weighed, and three samples were placed into 20 ml counting vials. Tissues were digested in 1 ml of 1N NaOH at 60°C overnight. The tissue digests were acidified by addition of 250 μ l

of 4N HCl prior to liquid scintillation counting (LSC). Plasma was isolated by centrifugation at 10,000 rpm for 5 minutes in a microfuge and duplicate 100 μ l aliquots of plasma were taken for LSC.

Cholesterol absorption, evaluated by this acute technique and expressed as the total amount of radioactive cholesterol in the plasma and liver, demonstrated that the wild type mice (+/+) absorbed an average of 11,773 dpm and *NPC1L1* deficient mice absorbed 992 dpm of the ¹⁴C-cholesterol. These results indicate that the *NPC1L1* deficient mice have a 92% reduction in cholesterol absorption. These data confirm the role of NPC1L1 in intestinal cholesterol absorption. Inhibition of NPC1L1-mediated cholesterol absorption, in a subject, by administering NPC1L1 antagonists, such as ezetimibe, to the subject, are a useful way to reduce serum cholesterol levels and the occurrence of atherosclerosis in the subject.

Example 24: Cholesterol Absorption in NPC1L1 (NPC3) Knockout Mice (Fecal Ratio Method: Cholesterol/Sitostanol).

In this example, cholesterol absorption and the activity of ezetimibe was determined in the *NPC1L1* knockout mice (-/-), heterozygous mice (+/-), and age matched wild-type mice (+/+).

Cholesterol absorption in the mice was determined by the dual fecal isotope ratio method as described by Altmann *et al.* (Biochim. Biophys. Acta. 1580(1): 77-93 (2002)). Mice (n= 4-6/group) were fed a standard rodent chow diet and in some groups treated daily with a maximally effective dose of ezetimibe (10 mg/kg). Mice were gavaged with ¹⁴C-cholesterol (1 μ Ci, 0.1mg unlabeled cholesterol) and ³H-sitostanol (2 μ Ci) in 0.1ml corn oil. Feces were collected for 2 days and fecal ¹⁴C-cholesterol and ³H-sitostanol levels were determined by combustion in a Packard Oxidizer. The fraction of cholesterol absorbed, as evaluated by the fecal dual isotope technique, was similar in wild type (+/+) and heterozygous mice (+/-) fed a chow diet (heterozygous mice absorbed 46 \pm 5% and age matched wild type mice absorbed 51 \pm 3% of the dose of ¹⁴C-cholesterol). The *NPC1L1* knockout mice (-/-) absorbed 15.6 \pm 0.4% of the ¹⁴C-cholesterol, which was similar to the wild type mice treated with a maximally effective dose of ezetimibe (16.1 \pm 0.3%), and reduced by 69% compared to wild type mice (p<0.001). In *NPC1L1* knockout treated with ezetimibe at 10 mg/kg/day, cholesterol absorption was similar to that seen in the untreated knockout mice (16.2 \pm 0.6% compared to 15.6% \pm 0.4%, respectively). Thus, the majority of

cholesterol absorption is dependent on the presence of NPC1L1 and the residual cholesterol absorption in mice lacking NPC1L1 is insensitive to ezetimibe treatment. These results indicate that NPC1L1 is involved in the small intestinal enterocyte uptake and absorption of cholesterol and is in the ezetimibe sensitive pathway.

5 **Example 25: Mouse Screening Assay (Acute Cholesterol Absorption).**

The following screening assay is used to identify the presence of an NPC1L1 antagonist in a sample.

Mice deficient in *NPC1L1* (-/-) are generated by breeding heterozygote mice (+/) to obtain wild-type (+/+) and *NPC1L1* deficient mice (-/-).

10 In a first set of experiments, non-fasted mice (6.5-9 weeks old, mixed 129 and C57BL/6 background) are weighed and grouped (n=1 to 4 -/- and n=1 to 4 +/+). All animals are gavaged (Feeding needles, 24G x 1 inch, Popper and Sons, NY) with 0.1 ml corn oil (Sigma; St. Louis, MO) containing 1 μ Ci ¹⁴C-cholesterol (New England Nuclear, [¹⁴C] Cholesterol, NEC-018) and 0.1mg carrier cholesterol mass
15 (Sigma; St. Louis, MO).

In another set of experiments, 1 to 4 wild-type *NPC1L1* mice (+/+) are treated identically to the mice in the first set of experiments, above, except that the mice are additionally fed a sample to be tested for the presence of an NPC1L1 antagonist.

20 Two hours later, blood is collected from each mouse by heart puncture. The liver is removed, weighed, and three samples are placed into 20 ml counting vials. Tissues are digested in 1 ml of 1N NaOH at 60°C overnight. The tissue digests are acidified by addition of 250 μ l of 4N HCl prior to liquid scintillation counting (LSC). Plasma is isolated by centrifugation at 10,000 rpm for 5 minutes in a
25 microfuge and duplicate 100 μ l aliquots of plasma are taken for LSC.

Cholesterol absorption, evaluated by this acute technique is expressed as the total amount of radioactive cholesterol in the plasma and liver. The sample tested is determined to contain an NPC1L1 antagonist when the level of cholesterol absorption (as measured by the above described methods) in the wild-type *NPC1L1*
30 mouse (+/+) which was fed the sample and in the *NPC1L1* deficient mouse (-/-) are less than the amount of cholesterol absorption in the wild-type *NPC1L1* mouse (+/+) which was not fed the sample.

**Example 26: Mouse Screening Assay
(Fecal Ratio Method: Cholesterol/Sitostanol).**

The following screening assay is used to identify the presence of an NPC1L1 antagonist in a sample.

5 Cholesterol absorption in the mice is determined by the dual fecal isotope ratio method as described by Altmann *et al.* (Biochim. Biophys. Acta. 1580(1): 77-93 (2002)).

Three groups of mice (n=1-6/group) are assembled. Two separate groups comprise wild-type *NPC1L1* mice (+/+) and one group comprises *NPC1L1* 10 deficient mice (-/-).

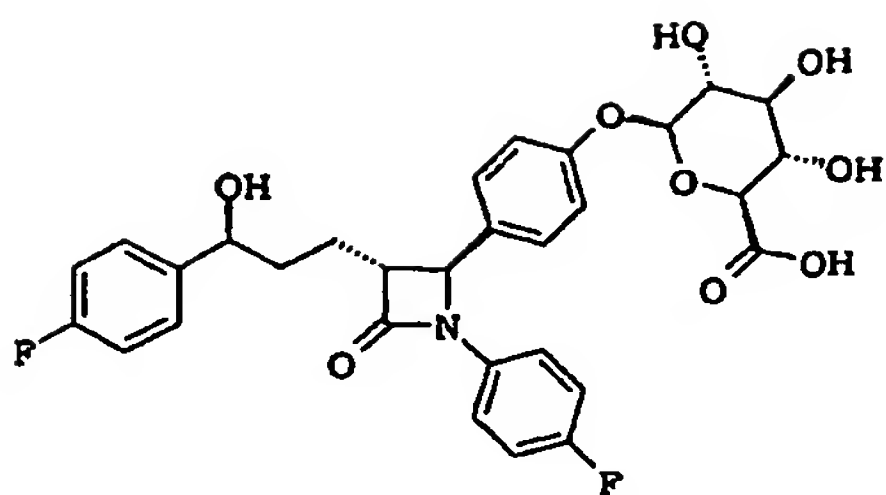
Each group is fed a standard rodent chow diet and in some groups treated daily. Mice are gavaged with ¹⁴C-cholesterol (1μCi, 0.1mg unlabeled cholesterol) and ³H-sitostanol (2μCi) in 0.1ml corn oil. One group of mice, which comprise wild-type *NPC1L1* mice (+/+) are further fed a sample to be tested for the 15 presence of an NPC1L1 antagonist. Feces are collected for 2 days and fecal ¹⁴C-cholesterol and ³H-sitostanol levels are determined by combustion in a Packard Oxidizer.

The sample tested is determined to contain an NPC1L1 antagonist when the level of cholesterol and/or sitostanol absorption (as measured by the above 20 described methods) in the wild-type *NPC1L1* mouse (+/+) which was fed the sample and in the *NPC1L1* deficient mouse (-/-) are less than the amount of cholesterol and/or sitostanol absorption in the wild-type *NPC1L1* mouse (+/+) which was not fed the sample.

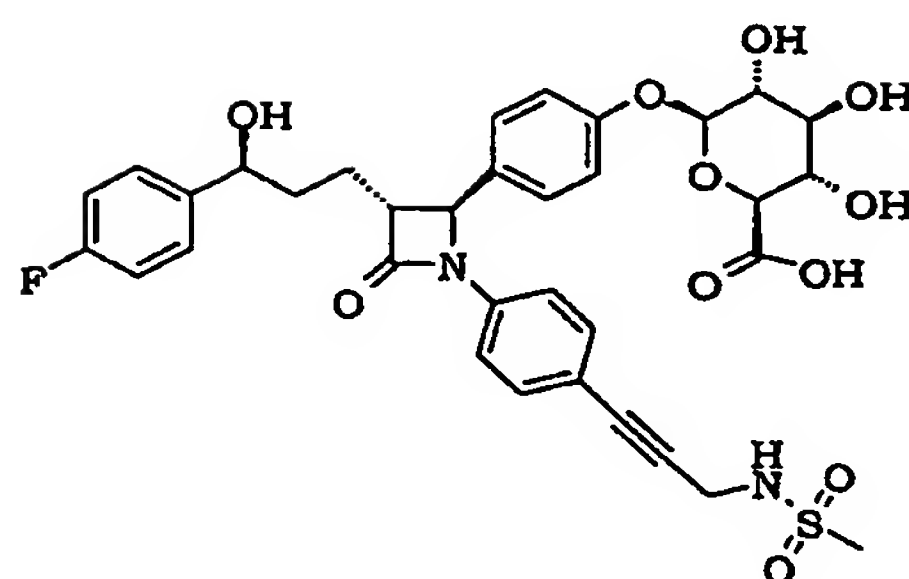
Example 27: Binding Analysis Using Brush Border Membrane Vesicles

25 The following screening assay may be used to identify the presence of an NPC1L1 ligand in a sample.

Materials. The following two compounds were synthesized for the binding assay described herein, ³H-ezetimibe glucuronide **1** (34.5 Ci/mmol) and its 30 ³⁵S-propargyl-sulfonamide derivative **2** (800-1100 Ci/mmol).

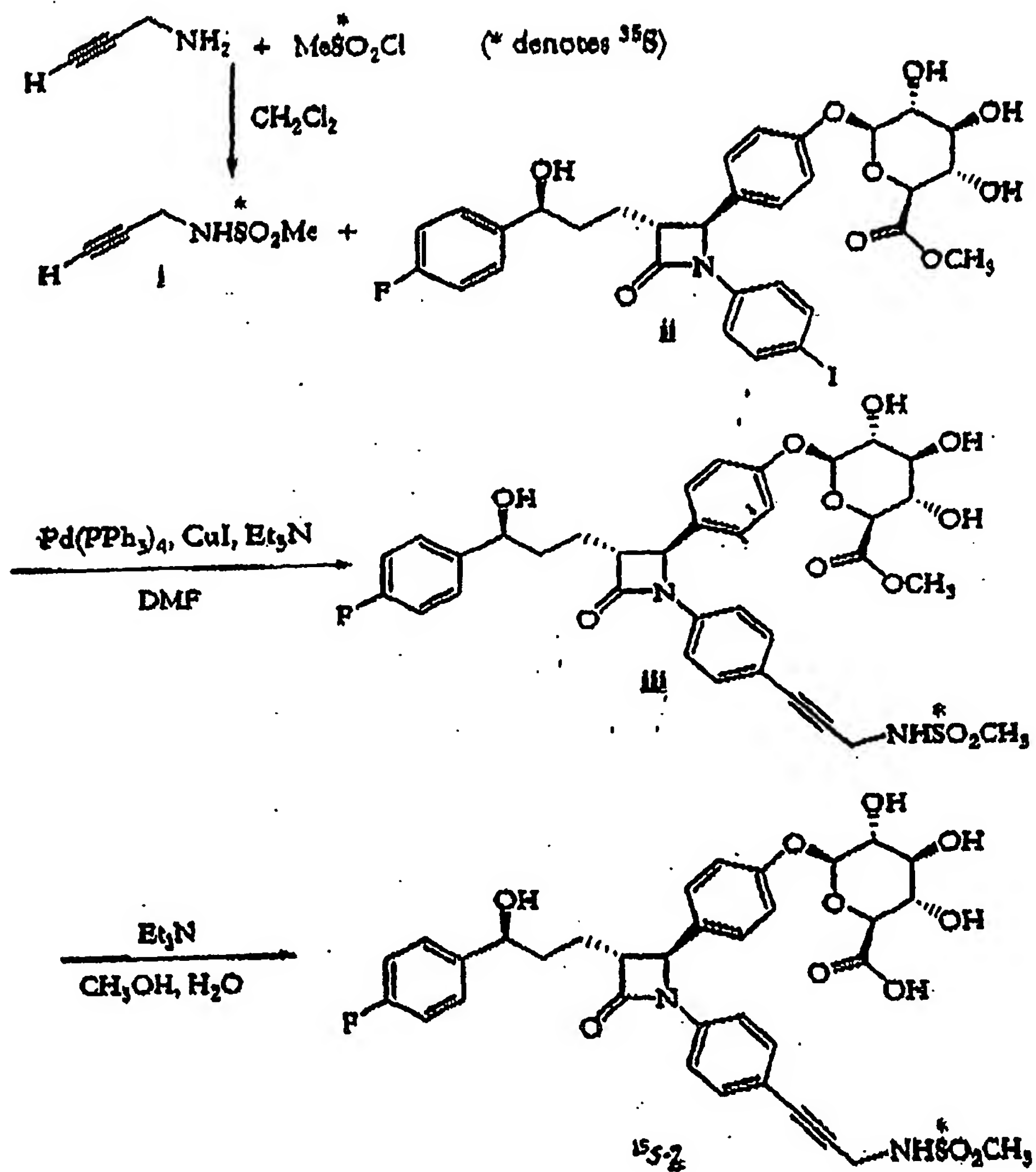


Ezetimibe-glucuronide
1



S-propargyl-sulfonamide
ezetimibe-glucuronide
2

Synthesis of ezetimibe glucuronide and S-propargyl-sulfonamide ezetimibe-glucuronide. Ezetimibe glucuronide (compound 1) (also referred to as EZE-glucuronide) can be made according to the procedures in U.S. Patent No. 5,756,470. The general scheme below illustrates a method for the synthesis of compound 2 and radiolabeled ^{35}S -2.



Preparation of compound ^{35}S -2 (Compound 2 with radiolabelled ^{35}S)

Step A: Preparation of [^{35}S]N-prop-2-yn-1-ylmethanesulfonamide (i).

The appropriate volume of [^{35}S]methane sulfonyl chloride (see Dean, D.C.; et al., *J. Med. Chem.* 1996, 39, 1767) totaling 3.5 mCi was removed from a stock solution
5 in methylene chloride and placed in a 5mL conical flask. It was then distilled at atmospheric pressure until the volume was approximately 50 μL . To this solution was immediately added 50 μL of propargylamine. After 15 min, the reaction mixture was diluted with 10 mL of ethyl acetate, washed with saturated sodium bicarbonate solution (3 x 2 mL), and dried over sodium sulfate. After filtration the resulting
10 solution had a count of 3.3 mCi and a radiochemical purity of 99.9 % by HPLC (Zorbax XDB C8 column, 4.6 x 150 mm, 5 % acetonitrile:H₂O (0.1 % TFA) to 100 % acetonitrile, 15 min linear gradient, 1 mL/min, t_R = 4.4 min).

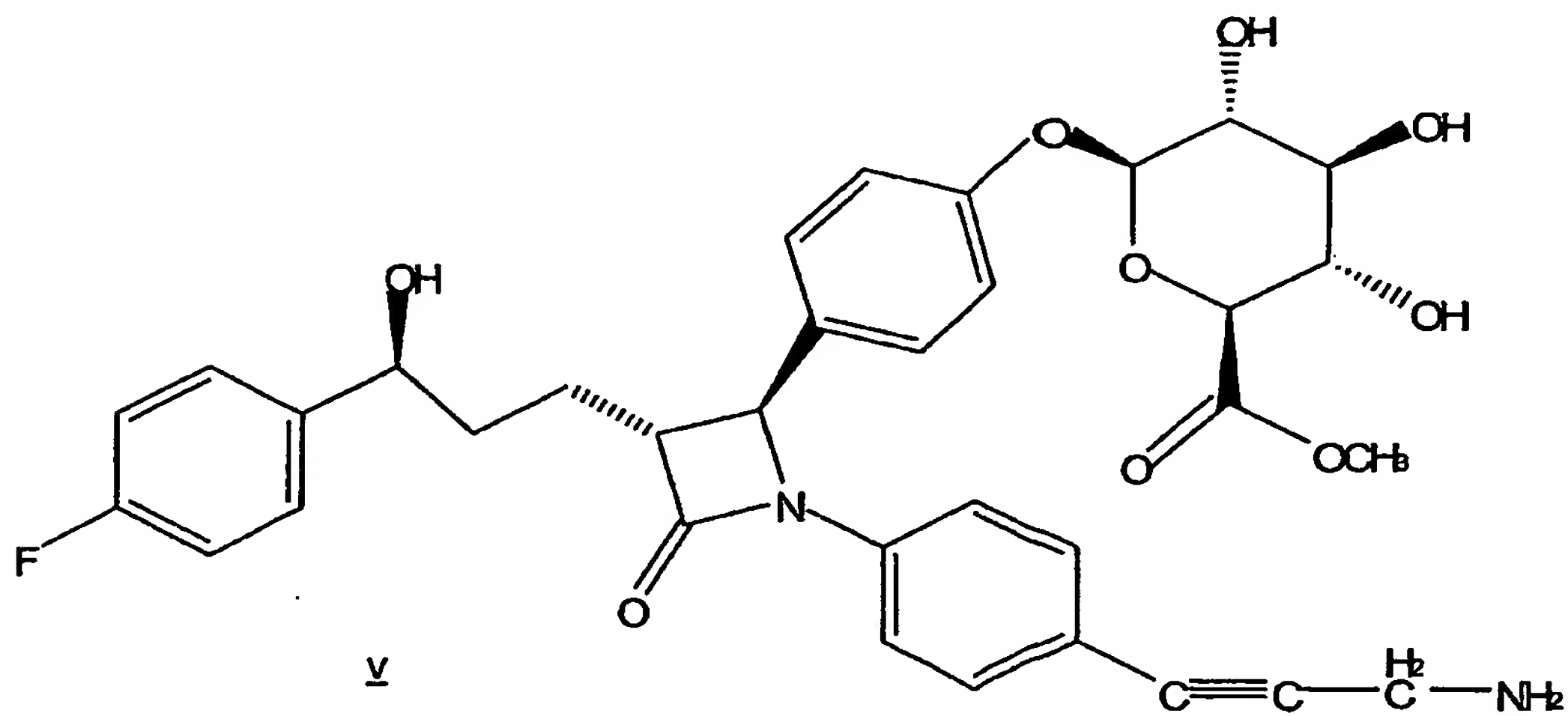
Step B: Preparation of [^{35}S] 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-{3-[(methylsulfonyl)amino]prop-1-yn-1-yl}phenyl)-4-oxoazetidin-2-yl]phenyl methyl β -D-glucopyranosiduronate ([^{35}S]) (iii). Dissolved
15 3.0 mCi of [^{35}S]N-prop-2-yn-1-ylmethanesulfonamide, 1 mg of compound ii (prepared according to Burnett, D.S. et al., *Bioorg. Med. Chem. Lett.* (2002), vol. 12, p. 311), and 1 μL of triethylamine in 100 μL of dimethylformamide inside a plastic microcentrifuge tube. To this solution was added 10 μL of a stock solution containing
20 8.1 mg of tetrakis(triphenylphosphine)palladium(0) and 1.4mg of copper iodide in 1 mL of dimethylformamide. Stirred at room temperature for sixty hours at which time HPLC indicated 55% conversion. This reaction mixture, which had a radiochemical purity of 44.4% by HPLC (Zorbax XDB C8 column, 4.6 x 150 mm, 5 % acetonitrile:H₂O (0.1 % TFA) to 100 % acetonitrile, 15 min linear gradient, 1
25 mL/min, t_R = 9.3 min) was taken on directly to the next step.

Step C: Preparation of [^{35}S] 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-{3-[(methylsulfonyl)amino]prop-1-yn-1-yl}phenyl)-4-oxoazetidin-2-yl]phenyl β -D-glucopyranosiduronic acid ^{35}S -2. The crude reaction mixture containing compound iii was treated with 25 μL of methanol, 90 μL of water,
30 and 30 μL of triethylamine and stirred at room temperature for one hour at which time it was concentrated to near dryness under a slow stream of nitrogen. The residue was dissolved in 1:1 acetonitrile:H₂O and subjected to semi-preparative chromatography

- (Zorbax XDB C8 250 x 9.4 mm column, 70:30 acetonitrile:H₂O (0.1 % TFA) 4 mL/min, 1 x 0.2 mL injections). 540 μ Ci of product was obtained which had a radiochemical purity of 99.9% by HPLC (Zorbax XDB C8 column, 4.6 x 150 mm, 70:30 acetonitrile:H₂O (0.1 % TFA), 1 mL/min, t_R = 10.4 min) and coeluted with an authentic sample of compound 2. LC/MS m/z = 508 (product – glucuronide – H₂O), SA = 769 Ci/mmol.

Alternate Preparation of ³⁵S-2.

- Step A: Preparation of iii. The appropriate volume of [³⁵S]methane sulfonyl chloride (see Dean, D.C.; et al., *J. Med. Chem.* 1996, 39, 1767) totaling 1.3 mCi was removed from a stock solution in methylene chloride and placed in a 5mL conical flask. It was then distilled at atmospheric pressure until the volume was approximately 50 μ L. To this solution was immediately added a solution of 1 mg of v in 5 μ L of pyridine (freshly distilled over calcium hydride).



15

- The solution was stirred at room temperature for five minutes at which time it was concentrated to near dryness under a slow stream of nitrogen. This reaction mixture, which had a radiochemical purity of 80.1% by HPLC (Zorbax XDB C8 column, 4.6 x 150 mm, 5% acetonitrile:H₂O (0.1 % TFA) to 100 % acetonitrile, 15 min linear gradient, 1 mL/min, t_R = 9.3 min) was taken on directly to the next step.

Step B: Preparation of ³⁵S-2. The crude reaction mixture containing iii was treated with 25 μ L of methanol, 90 μ L of water, and 30 μ L of triethylamine and

stirred at room temperature for one hour at which time it was concentrated to near dryness under a slow stream of nitrogen. The residue was dissolved in 1:1 acetonitrile:H₂O and subjected to semi-preparative chromatography (Zorbax XDB C8 250 x 9.4 mm column, 70:30 acetonitrile:H₂O (0.1 % TFA) 4 mL/min, 1 x 0.2 mL injections). 350 μ Ci of product was obtained which had a radiochemical purity of 98.4 % by HPLC (Zorbax XDB C8 column, 4.6 x 150 mm, 70:30 acetonitrile:H₂O (0.1 % TFA), 1 mL/min, t_R = 10.4 min) and coeluted with an authentic sample of 2. LC/MS m/z = 508 (product – glucuronide – H₂O), SA = 911 Ci/mmol.

Following the same general procedure for synthesis of ³⁵S-2, except omitting the radiolabel ling, compounds 2 and iv can be prepared.

Preparation of brush border membrane vesicles (BBMV).

Membranes were prepared from Rhesus macaque (*Macaca mulatta*), rat (male Sprague-Dawley), and mouse (male C57BL/6J) intestines, using Mg⁺⁺ precipitation method described in the following references and with modifications described below (Hauser, H., Howell, K., Dawson, R.M.C., Bowyer, D. E. *Biochim. Biophys. Acta* 602, 567-577 (1980); Kramer, W., Girbig, F., Gutjahr, U., Kowalewski, S., Jouvenal, K., Muller, G., Tripier, D., Wess, G. *J. Biol. Chem.* 268, 18035-18046 (1993); Rigrup, K.M., Ong, D.E. *Biochemistry* 31, 2920-2926 (1992)).

The intestines from freshly sacrificed animals were cut into segments, perfused with ice-cold saline buffer (Buffer A: 26 mM NaHCO₃, 0.96 mM NaH₂PO₄, 5 mM HEPES, 5.5 mM glucose, 117 mM NaCl, 5.4 mM KCl, pH = 7.4), placed on cold glass plates, opened longitudinally, and the mucosa scraped with glass microscope slips. This mucosa could be used fresh or frozen with identical results. To prepare the membranes, the mucosal scrapings were resuspended in 20 volumes of cold buffer consisting of 300 mM D-mannitol, 5 mM EGTA, 12 mM Tris, pH 7.4 with HCl, and containing 0.1 mM PMSF and a 1% dilution of a protease inhibitor cocktail (set 1, Calbiochem). They were homogenized using a Polytron at medium speed on ice until inspection with a microscope indicated complete cell lysis. Then, solid MgCl₂ was added slowly with stirring to a final concentration of 10 mM, and the solution was kept stirring on ice for 15 min. Cellular debris was removed by centrifugation for 15 min at 3,000g, and the membranes were recovered by centrifugation for 60 min at 48,000g. The membranes were further rinsed by re-suspension in a buffer containing 50 mM D-mannitol, 5 mM EGTA, and 2 mM Tris

at pH 7.40, and centrifugation for 60 min at 48,000g. The final pellet was resuspended in 120 mM NaCl and 20 mM Tris at pH 7.40 to a concentration of ~10-20 mg protein/ml, aliquoted, frozen in liquid nitrogen, and stored at -80C. The activity was stable indefinitely and could be freeze-thawed with minimal loss of activity.

Membrane protein was measured by the Bradford assay (Bradford, M.M. Anal. Biochem. 72, 248-254 (1976)) using bovine serum albumin as standard. The enrichment in brush border membranes was assessed using gamma-glutamyltransferase as a marker enzyme (Kramer, W., Girbig, F., Gutjahr, U., Kowalewski, S., Jouvenal, K., Muller, G., Tripier, D., Wess, G. J. Biol. Chem. 268, 18035-18046 (1993)), which indicated a 6-fold enrichment over the initial homogenate.

Binding assay. Assays were conducted in 12 x 75 mm glass test tubes and total volume 20-100 ul. In general, frozen membranes were diluted in buffer A or buffer A containing 0.03% taurocholate and 0.05% digitonin to a final concentration of 0.02 to 5 mg/ml. Radiolabeled ligands were typically 25-50 nM for ³H-ezetimibe (EZE)glucuronide 1, and 3-5 nM for its ³⁵S analog 2, in the assay, and they were delivered as DMSO or CH₃CN solutions. Competing ligands were likewise added as DMSO solutions to give a total 2-10 % organic solvent content. Nonspecific binding was defined by competition with 100 uM ezetimibe glucuronide. At least 2 components of buffer A, the bicarbonate and phosphate salts, were later found to be inconsequential and were routinely omitted. To ensure equilibrium was established, the reactions with compound 1 were incubated at least 3 hours for rhesus membranes and at least one hour for rat membranes at room temperature, and with compound 2 as long as 2 hours at 37°C with rhesus and rat brush membranes. Additionally, reactions with compound 2 were incubated as long as 2 hours at 37°C with membranes from HEK-293 cells expressing mouse, rat or human NPC1L1.

Bound ligand was quantified by single-tube vacuum filtration using GF/C glass fiber filters. Glass fiber filters (GF/C) were obtained from Whatman. The filters were pretreated by soaking with 0.5% polyethylenimine to reduce nonspecific binding. Filtration was accomplished by adding 2.5 ml of ice cold buffer (120 mM NaCl, 0.1% sodium cholate, and 20 mM MES at pH 6.70) to the assay tube, pouring the mixture through the filter, and then rinsing the tube and filter twice more with

another 2 x 2.5 ml buffer. The filters were counted in 7 ml vials using Packard DM liquid scintillation fluid. Where triplicate assays were performed, the standard error was typically <4%. As an example, a 100µl assay of rat brush border membranes at 2 mg/ml in the presence of 400,000 dpm (50 nM) ³H-ezetimibe glucuronide gave
5 15,000 dpm specific and 3,000 dpm nonspecific binding. The filters contributed most of the nonspecific binding (2,000 dpm).

Alternatively, vacuum filtration of compound 2 on a Millipore 96-well plate (Whatman GF/C) can also be used to achieve adequate precision.

Data analysis. Data from saturation experiments were subjected to a
10 Scatchard analysis, and linear regression was performed to yield the equilibrium dissociation constant (K_d) and maximum receptor concentration (B_{max}). Correlation coefficients for these determinations were typically greater than 0.96. Data from competition experiments were analyzed and IC_{50} values determined from Hill plots of the binding data. The kinetic data for ligand association and dissociation were
15 subjected to the analysis of Weiland and Molinoff (Weiland, G., Molinoff, .B. Life Sci. 29, 313-330 (1981)). The dissociation rate constant for (k_{-1}) was determined directly for a first order plot of ligand dissociation versus time. The rate of ligand association (k_1) was determined from the equation $k_1 = k_{obs}([LR]_e)/([L] [LR]_{max})$ where $[L]$ is the concentration of ligand, $[LR]_e$ is the concentration of the complex at
20 equilibrium, $[LR]_{max}$ is the maximum number of receptors present, and k_{obs} is the slope of the pseudo-first order plot $L_n ([LE]_e/([LR]_e - [LR]_t))$ versus time.

Binding analysis. Ezetimibe is rapidly converted to its glucuronide *in vivo*, and this metabolite is thought to be largely if not exclusively responsible for inhibition of cholesterol uptake. Accordingly, both ³H-ezetimibe and its
25 corresponding glucuronide derivative (1) were prepared and tested for binding to intestinal brush border membrane preparations, using a single-tube vacuum filtration technique. As a result of the hydrophobic nature of ³H-ezetimibe, high nonspecific binding was observed, precluding its use as a radioligand for the binding assay. However, due to the improved physical properties of the glucuronide derivative (1),
30 specific binding was observed with this radioligand and it was used to assess binding in rhesus, rat, and mouse intestinal brush border membranes.

Rhesus, rat, and mouse intestinal scrapings were homogenized and the brush border membranes isolated. Specific binding was observed exclusively in the

membrane fraction. Plots of total, nonspecific, and specific binding to rhesus (Figure 1) and rat (Figure 2) brush border membranes. Aliquots of rhesus BBMV (83 $\mu\text{g}/\text{assay}$) or rat BBMV (250 $\mu\text{g}/\text{assay}$) were incubated with increasing concentrations of ^3H -EZE-glucuronide. Total binding and nonspecific binding determined in the presence of 10-100 μM EZE-glucuronide are shown. Specific binding was calculated from the difference between total and nonspecific binding. Data were fit by nonlinear regression as described above, and the linear Scatchard plot is shown. In rhesus membranes, the data correspond to a single binding site with $K_d = 41 \text{ nM}$ and a concentration of 5.5 pmol/mg membrane protein. The affinity is ~ 10 -fold lower in rat membranes ($K_d = 540 \text{ nM}$). ^3H -EZE-glucuronide is not the best ligand for a binding assay for the mouse target due to the compounds low affinity in mouse membrane. These potencies correlate roughly with the sensitivity of these species to ezetimibe inhibition of cholesterol uptake.

Rate constants for binding and dissociation. Ezetimibe-glucuronide is slow to bind, and forms a relatively long-lived complex with its receptor. Indeed, this was key to detecting the interaction in a traditional filter-binding assay, as ligand/receptor interactions with K_d values greater than 100 nM often go unrecognized because of the typical fast off-rates of the ligands. Rate constants for association (k_{on}) and dissociation (k_{off}) of compound **1** were determined for rat and rhesus membranes, and used as an alternative method to calculate the dissociation constant (K_d) according to the relationship $K_d = k_{\text{off}}/k_{\text{on}}$. 300 $\mu\text{g}/\text{assay}$ of rat brush border membrane vesicles were incubated with 25 nM ^3H -EZE-glucuronide at room temperature for up to three hours for the association kinetic studies. 83 $\mu\text{g}/\text{assay}$ of rhesus brush border membrane vesicles were incubated with 25 nM ^3H -EZE-glucuronide at room temperature for up to five hours for the association kinetic studies. Nonspecific binding measured in the presence of 100 μM EZE-glucuronide was subtracted from the total binding to calculate the specific binding shown in Figures 3A and 4A. For the dissociation kinetic study, rat brush border membrane vesicles were incubated with 25 nM ^3H -EZE-glucuronide for 2 hours at room temperature and ligand dissociation was initiated by the addition of 100 μM EZE-glucuronide. Rhesus brush border membrane vesicles were incubated with 42 nM ^3H -EZE-glucuronide for 4 hours at room temperature and ligand dissociation was initiated by the addition of 100 μM EZE-glucuronide. For both rat and rhesus

dissociation studies, samples were collected at various times and radiolabel was detected. Dissociation curves are shown in Figures 3B (rat) and 4B (rhesus).

For rat membranes, the rate constant for association is $k_{on} = 5,540 \text{ M}^{-1} \text{ s}^{-1}$ (compared to 10^8 to $10^9 \text{ M}^{-1} \text{ s}^{-1}$ for diffusion controlled encounter), and the rate
5 constant for dissociation is $k_{off} = 2.4 \times 10^{-3} \text{ s}^{-1}$, corresponding to a half-life of 4.7 min. The data are shown in Figure 3, where the solid lines are theoretical for these rate constants. The K_d value predicted from these rate constants ($K_d = k_{off}/k_{on} = 440 \text{ nM}$) agrees well with that measured at equilibrium ($K_d = 540 \text{ nM}$).

For rhesus membranes, where ^3H -ezetimibe glucuronide is at least
10 10-fold more potent (as described above), the association rate remains the same but the half-life for dissociation of the complex increases to ~90 min. These data are shown in Figure 4, where the theoretical lines correspond to $k_{on} = 3,900 \text{ M}^{-1} \text{ s}^{-1}$ and $k_{off} = 1.23 \times 10^{-4} \text{ s}^{-1}$, and predict $K_d = 32 \text{ nM}$ compared to that measured at equilibrium ($K_d = 41 \text{ nM}$).

15 Example 28: Binding Analysis of a Potent NPC1L1 ligand

A ^{35}S -labeled propargyl-sulfonamide analogue of ezetimibe glucuronide (^{35}S -2) was identified as a potential NPC1L1 antagonist. Compound 2 was prepared as described in Example 27 and found to have markedly improved affinity for some species of brush border membranes vesicles. For rhesus brush
20 border membranes vesicles, 56 μg protein/assay were incubated with 25 nM ^3H -EZE-glucuronide in the presence of increasing concentrations of EZE-glucuronide and 2. For rat brush border membranes vesicles, 150 μg protein/assay were incubated with 50 nM ^3H -EZE-glucuronide in the presence of increasing concentrations of EZE-glucuronide and 2. For mouse brush border membranes vesicles, 20 μg protein/assay
25 were incubated with 3 nM ^{35}S -2 in the presence of increasing concentrations of EZE-glucuronide and 2.

2 is more potent against enterocyte brush border membrane preparations from rats (35-fold), but is equipotent with ezetimibe glucuronide for rhesus membrane preparations (Figure 5, Table 8). It also has enhanced affinity for
30 mouse membranes (Figure 6, Table 8).

Table 8. Summary of inhibition constants (K_i) for binding of ezetimibe glucuronide 1 and its propargyl-sulfonamide derivative 2 to rhesus, rat, and mouse intestinal brush border membranes.

Compound	Rhesus	Rat	Mouse
<u>1</u>	39	530	2,300
<u>2</u>	38	15	144

5 K_i values are nM.

Example 29: Distribution of ^3H -ezetimibe glucuronide (1) binding to intestinal tissues.

Previous studies have established that cholesterol absorption occurs primarily in the jejunum, and is substantially lower in the ileum and duodenum. To
 10 determine if the binding activity is similarly distributed, the binding assay using ^3H -ezetimibe glucuronide (^3H -1) as a radioligand was used to determine the distribution of binding sites in sections from rhesus and rat intestines.

For the rhesus studies, 10 cm corresponding to the ileum of a rhesus small intestine was separated and the remaining intestine was divided into three
 15 segments, (proximal, middle and distal) of equal length (70 cm each). For the rat studies, 10 cm corresponding to the ileum of a rhesus small intestine was separated and the remaining intestine was divided into three segments, (proximal, middle and distal) of equal length (36 cm each). Brush border membrane vesicles were prepared as described in Example 27. Aliquots of vesicles (100-200 μg) protein/assay were
 20 incubated with 50nM ^3H -EZE-glucuronide in the absence and presence of 100 μM EZE-glucuronide.

As shown in Figure 7, specific binding for ^3H -ezetimibe glucuronide peaks in the jejunum in both species, consistent with the previously observed pattern of cholesterol absorption.

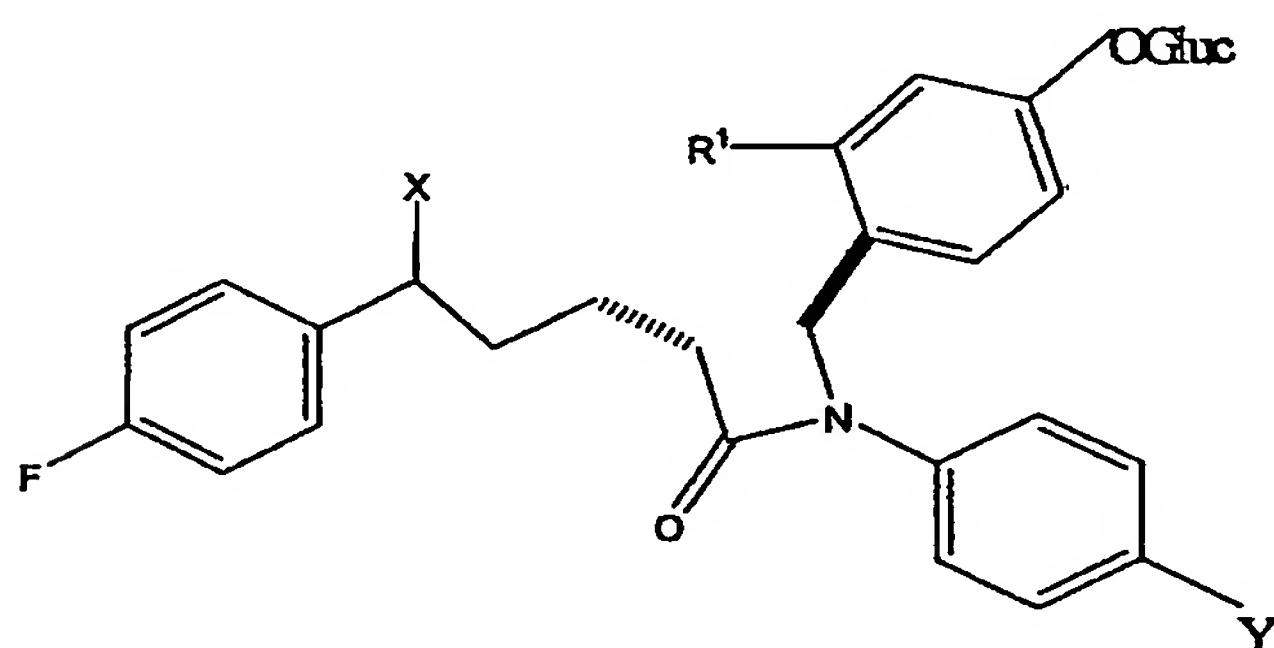
25 **Example 30: Correlation of *in vitro* and *in vivo* binding activity of NPC1L1**

To determine if *in vitro* binding activity is predictive of *in vivo* efficacy, the enantiomer of ezetimibe glucuronide and several close structural analogues of ezetimibe glucuronide that were tested in the rat membrane binding assay were tested in an acute rat cholesterol absorption study as described in
 30 Examples 23-26. The selected analogs had a range of *in vitro* potencies, and were

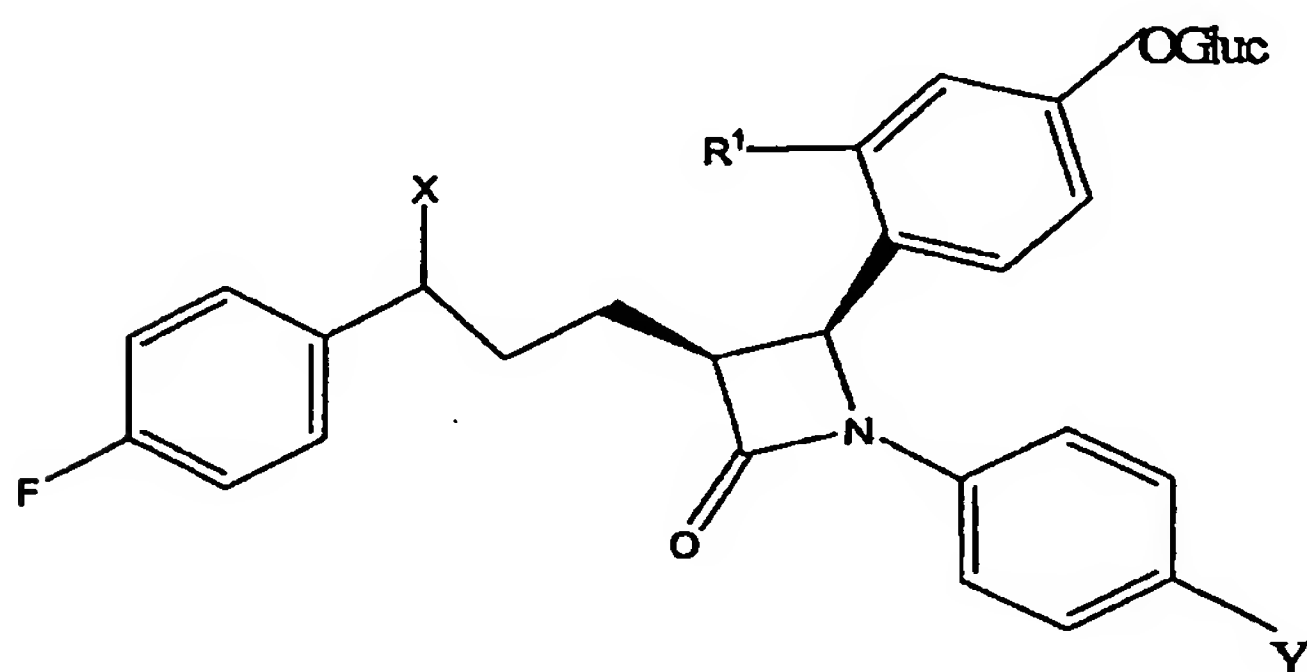
anticipated to have similar physical properties to ezetimibe glucuronide (Tables 9 and 12). The enantiomer, which has a $K_d > 100,000$ nM for the rat target, was inactive in the *in vivo* assay. For the other analogs, the same rank order of potency is observed in the *in vitro* and *in vivo* assays, further evidence that the observed binding is due to the target of ezetimibe.

Table 9. IC_{50} values of EZE-gluc and analogs to inhibit binding of 3H-EZE-gluc to rat brush border membrane vesicles.

Compound Name	R^1	X	Y	RAT IC_{50} (nM)
<u>3</u>	H	H2	F	2,300
EZE-gluc <u>1</u>	H	(S)-OH	F	530
EZE-gluc enantiomer <u>4</u>	H	(R)-OH	F	>100000
<u>5</u>	H	(R)-OH	F	3,900
<u>6</u>	H	=O	F	70,000
<u>7</u>	OH	(S)-OH	F	252



10 compounds 1, 3, 5, 6 and 7.



Backbone structure for compound 4.

Example 31: Binding affinities of ezetimibe glucuronide and its analogs to recombinant NPC1L1

NPC1L1 was identified as a candidate target of ezetimibe from a
5 search of genetic databases for cholesterol binding motifs. Subsequently, *NPC1L1*
deficient mice were found to have 80% reduction of cholesterol absorption, and did
not respond to ezetimibe treatment, strongly suggesting that this protein is required
for ezetimibe efficacy. To determine if NPC1L1 is the direct target of ezetimibe,
binding affinities were compared for ezetimibe glucuronide and several analogs to
10 NPC1L1 transfected cells and rat brush border membrane vesicles.

Rat NPC1L1 transfected CHO cells (~500,000 cells/assay) were
incubated with 5 nM ³⁵S-2 (~1 million dpm/assay) for 2 hours at 37 °C in the absence
or presence of increasing concentrations of EZE-glucuronide (compound 1),
compounds 2, 3, 5, 6, or 8. Compound 8 is an analog of compound 2 wherein the
15 hydroxyl group in the 3-hydroxylpropyl moiety of 2 is replaced with an oxo group.

Human NPC1L1 transfected CHO cells (~600,000 cells/assay) were
incubated with 5 nM ³⁵S-2 (~1 million dpm/assay) in buffer A for 2 hours at 37 °C in
the absence or presence of increasing concentrations of EZE-glucuronide (compound
1), compounds 2, 3, 5, 6, or 8.

20 As shown in Figures 9 and 12, and Table 10, the affinities for the
recombinant and native proteins are virtually identical, providing compelling evidence
that NPC1L1 is the direct target of ezetimibe in mammalian tissues, and that other
proteins are not required for binding.

Affinities of ezetimibe glucuronide and analogues thereof were also
25 determined for human recombinant NPC1L1. The results, shown in Figure 9, indicate
that ezetimibe glucuronide (1) has an affinity for the human protein of 907 nM. The
propargyl-sulfonamide analogue (2) is approximately 50-fold more potent, with a
 $K_d = 21$ nM, suggesting that this compound has the potential for enhanced potency of
cholesterol absorption inhibition in man.

Table 10: Comparison of inhibition constants (K_i) for binding to native rat intestinal brush border membranes and membranes from rat NPC1L1 transfected cells.

Analog	Recombinant rat NPC1L1 K_i , nM	Rat BBMV K_i , nM
EZE-glucuronide <u>1</u>	790	600
<u>2</u>	12	15
<u>3</u>	2400	2300
<u>6</u>	84500	70000
<u>5</u>	5800	3900
<u>8</u>	556	818

5 Example 32: Binding of ^{35}S -2 to membranes from wild type and NPC1L1 deficient mice.

Final confirmation that NPC1L1 is the target of ezetimibe was provided by binding studies with ^{35}S -2 in intestinal brush border membranes from NPC1L1 deficient and control mice.

10 Brush border membranes vesicles were prepared from intestinal tissues of wild type and *NPC1L1* knockout (-/-) mice. 15, 30 and 60 μg protein/ assay of brush border membranes vesicles were incubated with 4nM ^{35}S -2 in buffer A for 3 hours at 37°C in the presence and absence of 100 μM EZE-glucuronide.

15 The results, shown in Figure 10, indicate that no detectable binding is observed in membranes from NPC1L1 deficient mice, whereas age matched wild type control membranes have detectable binding. The binding affinity observed in this experiment in control mouse membranes ($K_d=156$ nM) was virtually identical to that observed in previous studies (Figure 11).

20 Example 33: Binding Analysis Using Brush Border Membrane Vesicles from Rat Mouse and Rhesus Monkey

Binding studies were performed to compare the relative binding affinity of ezetimibe glucuronide to various brush border membrane vesicles.

³H-ezetimibe glucuronide **1** was prepared as described in Example 27.

The brush border membranes were prepared as described in Example 27.

Binding Assay. Assays were conducted in 12 x 75 mm glass test tubes and total volume 20-100 μ l. In general, frozen membranes were diluted in buffer A or
5 buffer A containing 0.03% taurocholate and 0.05% digitonin to a final concentration of 0.5 to 5 mg/ml (Buffer A: 26 mM NaHCO₃, 0.96 mM NaH₂PO₄, 5 mM HEPES, 5.5 mM glucose, 117 mM NaCl, 5.4 mM KCl, pH = 7.4). Final concentrations of [³H]ezetimibe glucuronide **1** were typically 25-50 nM, and were delivered as DMSO or CH₃CN solutions. Competing ligands were likewise added as DMSO solutions to
10 give a total 1-5 % organic solvent content. Nonspecific binding was defined by competition with 100-500 μ M ezetimibe glucuronide. At least three components of buffer A, the bicarbonate and phosphate salts, and glucose, were later found to be inconsequential and were routinely omitted. Reactions were incubated until equilibrium was achieved (one hour for rat or three hours for rhesus membranes).

15 Bound ligand was recovered by single-tube vacuum filtration on Whatman GF/C glass fiber filters. The filters were pretreated by soaking with 0.5% polyethylenimine to reduce nonspecific binding. Filtration was accomplished by adding 2.5 ml of ice cold buffer (120 mM NaCl, 0.1 % sodium cholate, and 20 mM MES at pH 6.7) to the assay tube, pouring the mixture through the filter, and then
20 rinsing the tube and filter twice more with another 2 x 2.5 ml buffer. The filters were counted in 7 ml vials using Ultima Gold MV liquid scintillation fluid from Packard. Where triplicate assays were performed, the standard error was typically <4%. As an example, a 100 μ l assay of rat brush border membranes at 2 mg/ml in the presence of 400,000 dpm (50 nM) [³H]ezetimibe glucuronide gave 15,000 dpm specific and 3,000
25 dpm nonspecific binding. The filters contributed most of the nonspecific binding (2,000 dpm).

Data Analysis. After correction for nonspecific binding, saturation-binding data were fit by nonlinear regression (Sigma Plot) to the single-site expression $[B] = B_{\max} \times [L]/([L] + K_D)$. Linear Scatchard plots are shown for
30 illustration. Data on K_i from competition experiments were analyzed by nonlinear regression to the expression $[B] = [B_0]/(1 + [I]/K_i^{\text{obs}})$, and where required were corrected for radioligand competition as $K_i = K_i^{\text{obs}}/(1 + [L^*]/K_D)$.

First-order rate constants (k_{obs} and k_{off}) were determined by nonlinear regression to the first order rate equation $A = A_0 e^{-kt}$. Kinetic data for k_{on} were analyzed according to Weiland and Molinoff (32), using the equation $k_{\text{on}} = k_{\text{obs}} ([LR]_e / ([L] [LR]_{\text{max}}))$, where $[L]$ is the concentration of ligand, $[LR]_e$ is the concentration of the complex at equilibrium, $[LR]_{\text{max}}$ is the maximum number of receptors present, and k_{obs} is the apparent first-order rate constant.

Binding analysis. Binding studies using the [^3H]ezetimibe glucuronide a traditional rapid-filtration assay on glass fiber filters using enterocyte brush border membrane preparations from rat, mouse and rhesus monkey were performed (Table 11). Table 11 shows the binding affinities of [^3H]ezetimibe glucuronide to the membranes in the absence of detergents. The observed binding was specific, saturable, and consistent with a single molecular site. Scatchard analyses and the specific/nonspecific binding windows for rat and monkey are shown in Figure 12. The binding affinity is relatively weak in rat membranes ($K_D = 542$ nM) and even weaker in murine membranes ($K_D = 10,000$ nM). In contrast, binding affinity in rhesus monkey membranes is approximately 10-fold greater ($K_D = 41$ nM). The number of binding sites varied from 5 – 20 pmol/mg membrane protein depending on species and preparation.

The rates for binding and dissociation of [^3H]ezetimibe glucuronide were determined and found to be slow relative to those typically observed for protein-ligand interactions. For example, the rate constants for association to rat and monkey brush border membranes are $k_{\text{on}} = 5.54$ and $3.90 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ (Figure 12). These are 100,000-fold smaller than those typically observed for a diffusion controlled encounter, 10^8 to $10^9 \text{ M}^{-1} \text{ s}^{-1}$. Similarly, these complexes are unusually long-lived, dissociating with rate constants of $k_{\text{off}} = 2.4 \times 10^{-3} \text{ s}^{-1}$ and $1.2 \times 10^{-4} \text{ s}^{-1}$ at 25C, equivalent to half-lives of 4.7 and 96 min for the rat and monkey complexes, respectively. In comparison, half lives are normally <1 sec for dissociation of common diffusion controlled, 100 nanopolar K_D ligands. These rate constants predict K_D values ($K_D = k_{\text{off}}/k_{\text{on}}$) of 440 and 32 nM, respectively, which agree well with those measured by equilibrium titration (Figure 12), and by saturation as described earlier. Such slow-forming, long-lived complexes suggest that conformational changes in the protein are rate limiting.

Table 11: Comparison of ezetimibe binding affinity and cross species efficacy

Species	K _D (nM)	ED ₅₀ (mg/Kg)
mouse	12,000	0.5000
rat	542	0.0300
monkey	41	0.0005

Table 11 also shows a correlation between *in vitro* and *in vivo* binding of [³H]ezetimibe glucuronide in various enterocyte brush border membrane preparations from rat, mouse and rhesus monkey. The *in vivo* ED₅₀ values are derived from cholesterol absorption and cholesterol feeding studies. The rank order of ezetimibe potency (ED₅₀) *in vivo* as follows: rhesus (0.0005mpk) > rat (0.03mpk) > mouse (0.5mpk) is the same as the order of *in vitro* binding affinity (IC₅₀) as follows: rhesus monkey (41nM) < rat (542nM) < mouse (12,000nM).

The binding affinities of **1** to brush border membranes correlate well across species with the sensitivity to ezetimibe inhibition of cholesterol uptake *in vivo* (mouse < rat < monkey) (Clader, J. W. The discovery of ezetimibe; A view from outside the receptor. *J Med. Chem.* 47, 1-9 (2004); Davis, H.R. Jr., Compton, D.S., Hoos, L. & Tetzloff, G. Ezetimibe, a potent cholesterol absorption inhibitor, inhibits the development of atherosclerosis in ApoE knockout mice. *Arterioscler. Thromb. Vasc. Biol.* 21, 2032-2038 (2001); Burnett, D.A. Beta-lactam cholesterol absorption inhibitors. *Curr. Med. Chem.* 11, 1873-1887 (2004), consistent with the hypothesis that the assay is relevant to the target of ezetimibe *in vivo* (Table 11). As evidence that this interaction is very specific, the glucuronide of the enantiomer of ezetimibe was prepared and found to be completely inactive *in vitro* (K_i > 100 x K_D for ezetimibe glucuronide in all species), consistent with its lack of activity *in vivo* in a rat acute cholesterol absorption model (see Table 12 in which the enantiomer is analyzed).

Example 34: NPC1L1 as the target of ezetimibe in NPC1L1-expressing HEK293 cells.

This example demonstrates that ezetimibe binds specifically to NPC1L1-expressing HEK293 cells.

Transient expression of NPC1L1. Plasmid pCR3.1 expressing rat NPC1L1 (Genbank AY437867) or human NPC1L1 (Genbank AY437865) were

prepared using standard molecular biology protocols. HEK-293 cells (ATCC) were seeded at 10×10^6 cells per T-225 flask (Corning) in DMEM containing 10% fetal calf serum, 4.5 g/L D-glucose and L-glutamine, 18 hours prior to transfection. They were transiently transfected with 25 μ g of DNA using Fugene transfection reagent (Roche Biochemical) at a ratio of 6:1 Fugene:DNA. Following transfection, the cells were incubated at 37°C and 5% CO₂ for 48 hours, and then harvested using PBS based cell dissociation buffer (Gibco), pelleted at 500 x g, snap frozen on dry ice, and stored at -80°C.

Membrane preparation from HEK-293 cells. Membranes were prepared by resuspending the frozen cell pellets in ten volumes of 20 mM HEPES/Tris buffer at pH 7.40 containing 8% sucrose, and sonicating the suspensions with a probe sonicator on ice until most of the cells were lysed. To isolate the membranes, the sonicates were centrifuged at 1600 x g for 10 min to remove cell debris, and then the supernatants were centrifuged at 125,000 x g for 1 hour to recover the membranes. These membranes were resuspended in 20 mM HEPES/Tris buffer at pH 7.40 containing 160 mM NaCl and 5% glycerol, and stored at 10-20 mg/ml protein at -80°C.

Pursuing the recent evidence indicating that NPC1L1 is an important component of the pathway inhibited by ezetimibe, recombinant rat and human NPC1L1 were expressed in human embryonic kidney (HEK) 293 cells (Figure 13, Panel 1). Cell lysates from HEK-293 cells expressing NPC1L1 (Lanes 1 and 3 of Panel 1 Figure 13) and wild-type cells (Lanes 2 and 4 of Panel 1 Figure 13) were analyzed by gel electrophoresis and Western blot with an anti-NPC1L1 antibody A1801. An excess of NPC1L1-specific peptide was included to assess specificity of the antibody for NPC1L1 (Lane 3 and 4 of Panel 1 Figure 13). Preliminary binding studies using 1 revealed specific binding to membrane preparations from cells expressing NPC1L1, and no specific binding to membranes from mock transfected cells (not shown).

Binding to NPC1L1 expressing cells was also observed with a BODIPY-labeled fluorescent ezetimibe glucuronide analog (SCH354909) (Figure 13, Panel 2A). Panel 2 of Figure 13 shows confocal microscope images of a fluorescent ezetimibe glucuronide analog (SCH354909) bound to the surface of NPC1L1-293 cells (Panel 2A), nonspecific binding of SCH354909 to NPC1L1-293 cells in the

presence of 100 μ M unlabeled ezetimibe glucuronide (Panel 2B), binding of SCH354909 to wild type HEK 293 cells (Panel 2C), and nonspecific binding of SCH354909 to wild type HEK 293 cells in the presence of 100 μ M unlabeled ezetimibe glucuronide (Panel 2D). In each case, plated cells were incubated in culture media with 500 nM SCH354909 for 4 hours at 37°C. Cells were subsequently washed with PBS and fluorescence was detected using confocal microscopy.

Binding of SCH354909 was clearly evident at the cell surface membrane of the NPC1L1-expressing cells and was completely abolished in the presence of excess unlabeled ezetimibe glucuronide (Figure 13, Panel 2C). No binding was observed in wild type HEK 293 cells (Figure 13, Panels 2B and 2D). These results demonstrated that ezetimibe glucuronide binds specifically to NPC1L1.

Example 35: NPC1L1 as the *in vivo* target of ezetimibe.

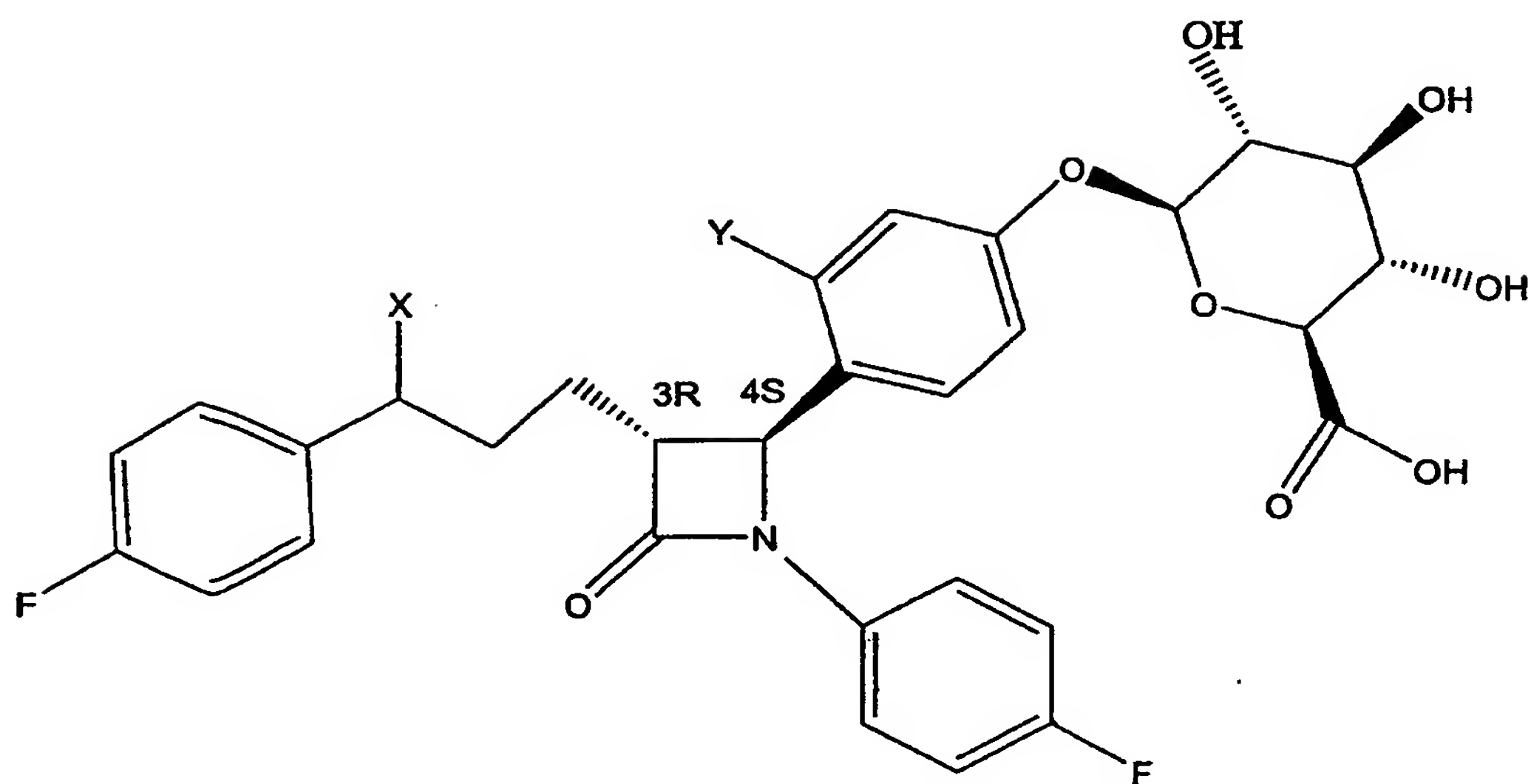
To obtain evidence that NPC1L1 is the direct binding target of ezetimibe *in vivo*, binding affinities of 1 and several key analogs were determined for recombinant rat and human NPC1L1 expressed in HEK-293 cell membranes and compared to those for native rat and rhesus intestinal enterocyte brush border membranes. A series of ezetimibe analogs was selected with subtle structural diversity, but with no binding affinities to native brush border membranes that covered a range of 1000-fold.

Table 12 shows a comparison of binding affinities (K_i values) for recombinant NPC1L1-293 cell membranes and native brush border membranes. Selected analogs of ezetimibe glucuronide are compared against recombinant rat and human NPC1L1 membranes prepared from transiently transfected HEK-293 cells compared to native rat and rhesus brush border membranes. The binding assays were conducted in a final volume of 20 μ l in the presence of 0.03% sodium taurocholate and 0.05% digitonin until equilibrium was achieved. 1.25 mg protein/ml and 100 nM 1 were used for native rat, recombinant rat, and recombinant human experiments, and 1.25 mg protein/ml and 20 nM 1 were used for native rhesus monkey experiments. Observed total and nonspecific binding, respectively, in the absence of inhibition were native rat: 7,700 & 1,100, recombinant rat: 33,000 & 1,100, native rhesus monkey: 7,300 & 367, and recombinant human: 19,200 & 1,000 dpm. Analog structures are defined in Table 12. Compound 4 has the stereochemical configuration 3S, 4R, and is the glucuronide of the enantiomer of ezetimibe. These determinations

were conducted in buffer containing 0.03% taurocholate and 0.05% digitonin, levels below the critical micelle concentrations of these detergents. These conditions enhanced apparent binding by as much as 20-fold for the recombinant preparations (principally a B_{max} effect), and greatly facilitated a quantitative comparison of K_i values for 1 and its analogs.

As shown in Table 12, the K_i values for recombinant rat NPC1L1 and native rat brush border membranes are virtually identical, strongly suggesting that NPC1L1 is the molecular target of ezetimibe *in vivo*. In the case of membranes from cells expressing recombinant human NPC1L1, the binding affinities also parallel those observed in rat membranes, whereas binding affinities for native rhesus brush border membranes are uniformly ~10-fold more potent. This result is consistent with the finding that ezetimibe is an order of magnitude more potent in monkey than in human or rat (Clader, J. W. The discovery of ezetimibe; A view from outside the receptor. *J Med. Chem.* 47, 1-9 (2004); Jeu, L. & Cheng, J.W. Pharmacology and therapeutics of ezetimibe (SCH 58235), a cholesterol-absorption inhibitor. *Clin. Ther.* 25, 2352-2387 (2003)).

TABLE 12



5

Analog	X	Y	Rat BBM Ki (nM)	Rat NPC1L1 Ki (nM)	Human NPC1L1 Ki (nM)	Monkey BBM Ki (nM)
EZE-gluc <u>1</u>	OH (<i>S</i>)	H	390	210	220	15
<u>4</u> *	OH (<i>R</i>)	H	130,000	74,000	130,000	22,000
<u>3</u>	H	H	1,600	820	1,000	150
<u>6</u>	=O	H	33,000	23,000	14,000	3,300
<u>5</u>	OH (<i>R</i>)	H	2,800	1,700	1,300	120
<u>7</u>	OH (<i>S</i>)	OH	280	360	210	60

* glucuronide of the enantiomer of ezetimibe has stereochemical configuration 3*S*, 4*R*

Conclusive evidence that NPC1L1 is the target of ezetimibe was provided by studies with tissues from NPC1L1 deficient mice. Enterocyte brush border membranes prepared from NPC1L1 deficient mice showed no detectable specific binding affinity for 1, whereas membranes from age-matched wild-type mice showed a high level of specific binding with a $K_D = 12 \mu M$ (Figure 14).

For Figure 14A, enterocyte brush border membranes were prepared from NPC1L1 deficient male mice and same sex wild-type littermates, and tested for binding of 1. Conditions for binding were 5 mg/ml protein and 500 nM 1 in a volume of 20 μl and in the presence of 0.03% sodium taurocholate and 0.05% digitonin. Membranes from wild type mice are on the left and from NPC1L1 deficient mice on the right. The bar graphs indicate total binding (left bar), nonspecific binding in the presence of 500 μM cold ezetimibe glucuronide (middle bar), and specific (right bar) binding, respectively for each of wild type and NPC1L1 deficient mice, and error bars represent triplicate measurements. The graphs show that although specific binding is readily detectable in wild-type mice, it is absent in NPC1L1-deficient mice.

Figure 14B shows a plot demonstrating competition of unlabeled ezetimibe glucuronide against 1. Membranes from wild-type mice (upper curve) gave $K_i = 12,000$ nM, while specific binding was virtually undetectable in membranes from the knockout animals (lower curve). Conditions were those described in Figure 14A.

The present studies involve a quantitative comparison of binding between recombinant proteins and brush border membranes. SR-BI (scavenger receptor type B1) was previously identified as a potential target using an expression cloning strategy employing ezetimibe binding to candidate proteins; this hypothesis was readily dismissed when neither cholesterol absorption nor ezetimibe activity were affected in SR-BI deficient mice (Altmann, S. W. *et al.* The identification of intestinal scavenger receptor class B, type 1 (SR-B1) by expression cloning and its role in cholesterol absorption. *Biochem. Biophys. Acta* 1580, 77-93 (2002)). The results show that ezetimibe binds to native intestinal membranes and cells expressing recombinant NPC1L1 with comparable affinity, and does not bind to membranes from NPC1L1 deficient mice, indicating a specific binding interaction between NPC1L1 and ezetimibe. Together with the previously published findings that mice deficient in NPC1L1 are defective in intestinal cholesterol uptake, and are no longer responsive to ezetimibe (Altmann, S. W. *et al.* Niemann-Pick C1 Like 1 protein is critical for

intestinal cholesterol absorption. *Science* 303, 1201-1204 (2004)), these data definitively establish NPC1L1 as the direct target of ezetimibe.

Example 36: Effect of detergents on [³H]ezetimibe glucuronide binding.

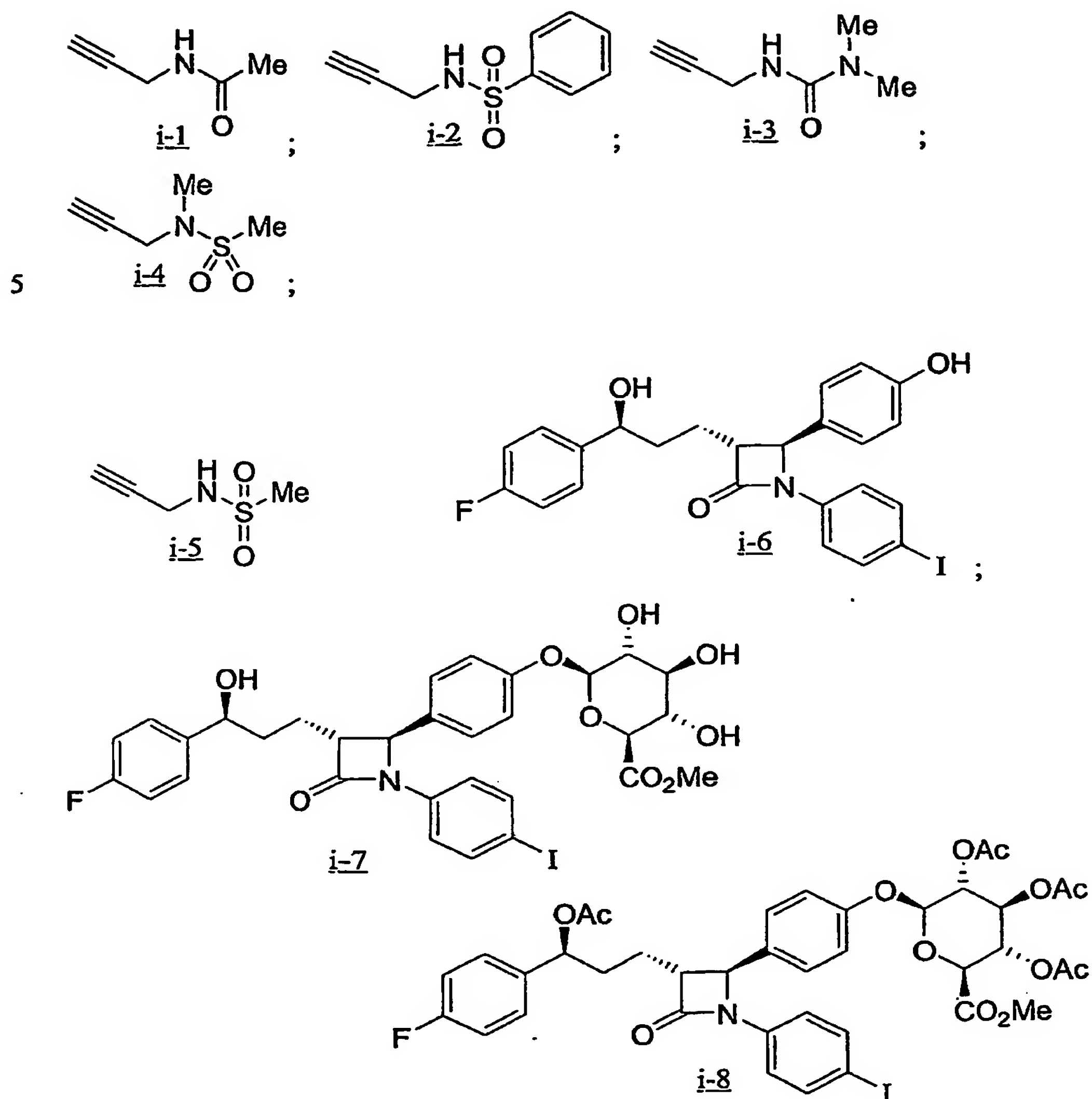
A practical aspect of work with the recombinant protein was that the number of binding sites in transfected NPC1L1-293 cell membranes initially appeared quite low. The influence of a combination of 0.03% taurocholate and 0.05% digitonin on specific binding to these and native enterocyte brush border membrane preparations is dramatic as illustrated in Figure 15.

Equal amounts (25 µg protein) of rat brush border membranes, membranes from HEK-293 cells transiently expressing recombinant rat and human NPC1L1, were incubated with 25 nM 1 in a final volume of 20 µl until equilibrium was achieved. The incubation conditions were buffer A with and without sodium taurocholate and digitonin to a final concentration of 0.03% and 0.05%, respectively. On the x-axis, "C" denotes controls in the absence of detergent, and "+det" the response in the presence of both detergents. The results are shown in 3 bar groupings; Total binding (left bar in each 3 bar group), nonspecific binding in the presence of 100 µM unlabeled ezetimibe glucuronide (middle bar in each 3 bar group), and specific binding (right bar in each 3 bar group) are shown.

Example 37: Binding affinities of ezetimibe glucuronide and various analogues NPC1L1 in rat and rhesus monkey membranes.

As determined from binding assay results using ³H-ezetimibe glucuronide with rat brush border membrane, representative tested compounds of Formula II were determined to have IC₅₀'s of about 13,000nM or lower, and particularly certain tested compounds had IC₅₀'s of about 1900nM or lower, more particularly certain tested compounds had IC₅₀'s of about 1000nM or lower, and most particularly certain tested compounds had IC₅₀'s of less than 100nM. As determined from binding assay results using ³H-ezetimibe glucuronide with rhesus brush border membrane, representative tested compounds of Formula II were determined to have IC₅₀'s of about 4200nM or lower, and particularly certain tested compounds had IC₅₀'s of about 165nM or lower, more particularly certain tested compounds had IC₅₀'s of less than 100nM, and most particularly certain tested compounds had IC₅₀'s of less than 50nM.

The designations below are used in the Examples that follow for certain repetitively used intermediates:

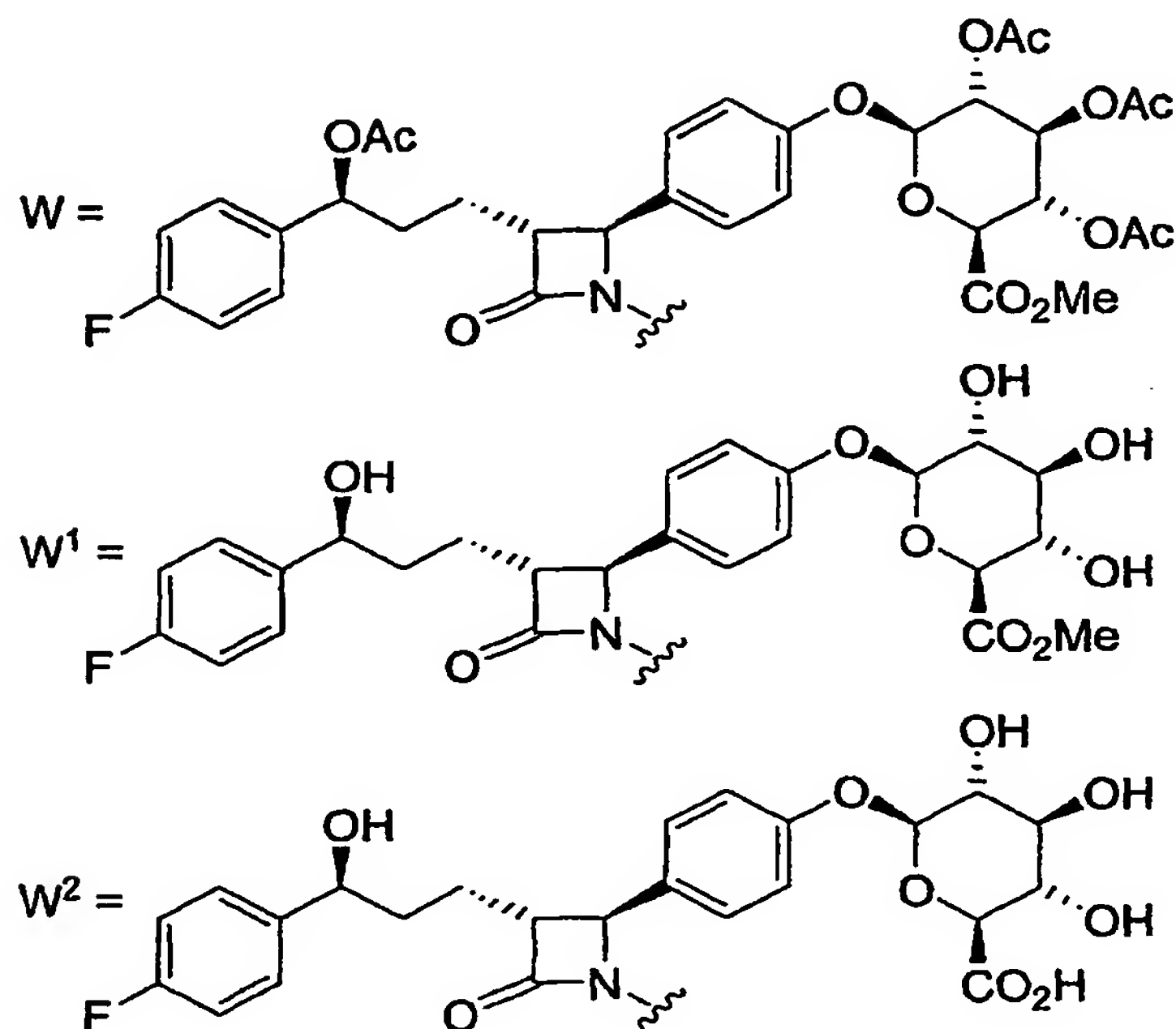


The compounds (3*R*,4*S*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-4-(4-hydroxyphenyl)-1-(4-iodophenyl)azetidin-2-one (i-6) and 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-iodophenyl)-4-oxoazetidin-2-yl]phenyl methyl β-D-glucopyranosiduronate (i-7) were prepared according to Burnett, D. S.; Caplen, M. A.; Domalski, M. S.; Browne, M. E.; Davis, H. R. Jr.; Clader, J. W. *Bioorg. Med. Chem. Lett.* (2002), 12, 311. Compound i-8 is the hydroxy-protected analog of i-7, where the protecting group is acyl.

10

15

The following definitions are also used in the Examples that follow:



EXAMPLE 38

5 Preparation of *N*-prop-2-yn-1-ylacetamide (i-1)

Acetyl chloride (0.52 mL, 7.3 mmol) was added to a stirred solution of propargylamine (0.5 mL, 7.3 mmol) and dimethylaminopyridine (18 mg, 0.14 mmol) in pyridine (2.5 mL) at 0°C, and the resulting mixture was allowed to warm to ambient temperature. After approximately 15 h, the reaction mixture was diluted with ethyl acetate and washed successively with 1N HCl and brine. The organic phase was dried (Na₂SO₄), filtered and concentrated *in vacuo* to afford the title compound (i-1), which was used without further purification.

EXAMPLE 39

Preparation of *N*-prop-2-yn-1-ylbenzenesulfonamide (i-2)

15 Benzene sulfonyl chloride (1.16 mL, 9.1 mmol) was added to stirred solution of propargylamine (0.62 mL, 9.1 mmol) and dimethylaminopyridine (22 mg, 0.18 mmol) in pyridine (5 mL) at room temperature. The resulting solution was aged

at ambient temperature for approximately 15 h. The reaction mixture was diluted with ethyl acetate and washed successively with 1N HCl and brine. The organic phase was dried (Na₂SO₄), filtered and concentrated *in vacuo* to furnish the title compound (i-2), which was used without further purification.

5

EXAMPLE 40

Preparation of *N,N*-Dimethyl-*N'*-prop-2-yn-1-ylurea (i-3)

Dimethyl carbamylchloride (0.84 mL, 9.1 mmol) was added to a stirred solution of propargylamine (0.62 mL, 9.1 mmol) and dimethylaminopyridine (22 mg, 0.18 mmol) in pyridine (5 mL) at room temperature. The resulting
10 suspension was stirred at ambient temperature for approximately 15 h. The reaction mixture was diluted with ethyl acetate and washed successively with 1N HCl and brine. The organic phase was dried (Na₂SO₄), filtered and concentrated *in vacuo* to afford the title compound (i-3), which was used without further purification.

EXAMPLE 41

15 Preparation of *N*-Methyl-*N*-prop-2-yn-1-ylmethanesulfonamide (i-4)

Methansulfonylchloride (1.12 mL, 14.5 mmol) was added to a stirred solution of *N*-methylpropargylamine (1.22 mL, 14.5 mmol) and dimethylaminopyridine (35 mg, 0.30 mmol) in pyridine (10 mL) at room temperature. After aging for approximately 15 h, the reaction mixture was poured into ethyl acetate
20 and washed successively with 1N HCl and brine. The organic phase was dried (Na₂SO₄), filtered and concentrated *in vacuo*, to afford the title compound (i-4), which was used without further purification.

EXAMPLE 42

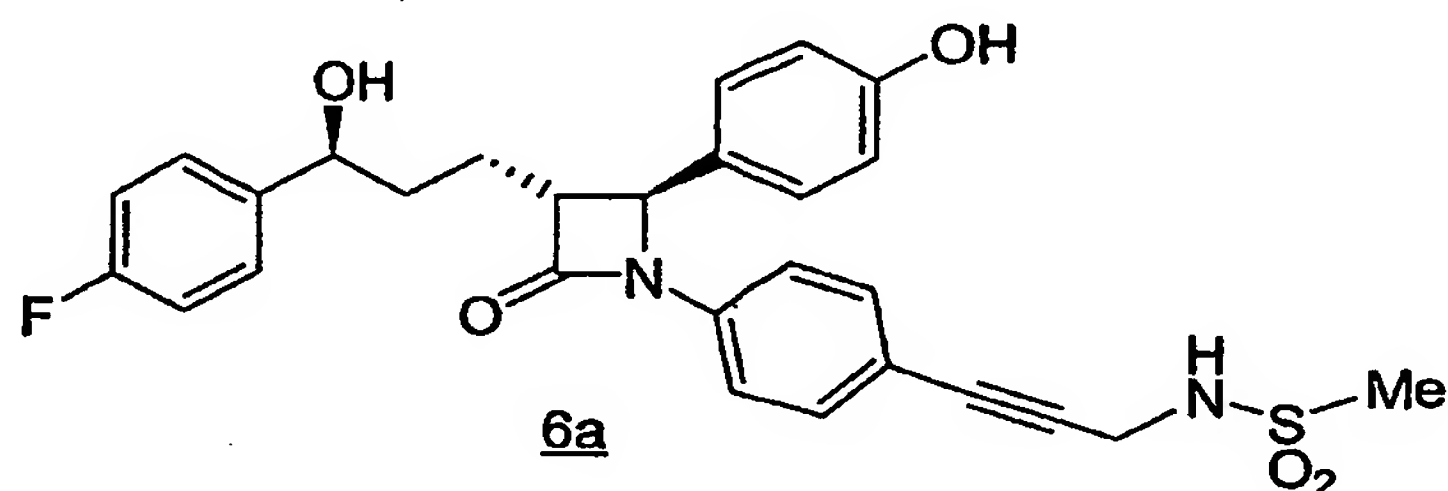
Preparation of *N*-prop-2-yn-1-ylmethanesulfonamide (i-5)

25 Methansulfonylchloride (1.40 mL, 18.1 mmol) was added dropwise to a stirred solution of propargylamine (1.00 g, 18.1 mmol) and dimethylaminopyridine (44.0 mg, 0.36 mmol) in pyridine (10 mL) at 0 °C. After aging for approximately 15 h, the reaction mixture was poured into 1N HCl and extracted twice with ethyl acetate. The combined organic extracts were washed with saturated aqueous sodium
30 bicarbonate, brine, dried (MgSO₄), filtered and concentrated *in vacuo*, to afford the

title compound i-5. Crude i-5 crystallized on standing and was used without further purification.

EXAMPLE 43

Preparation of *N*-(3-{4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-2-(4-hydroxyphenyl)-4-oxoazetidin-1-yl]phenyl}prop-2-yn-1-yl)methanesulfonamide (Compound 6a)

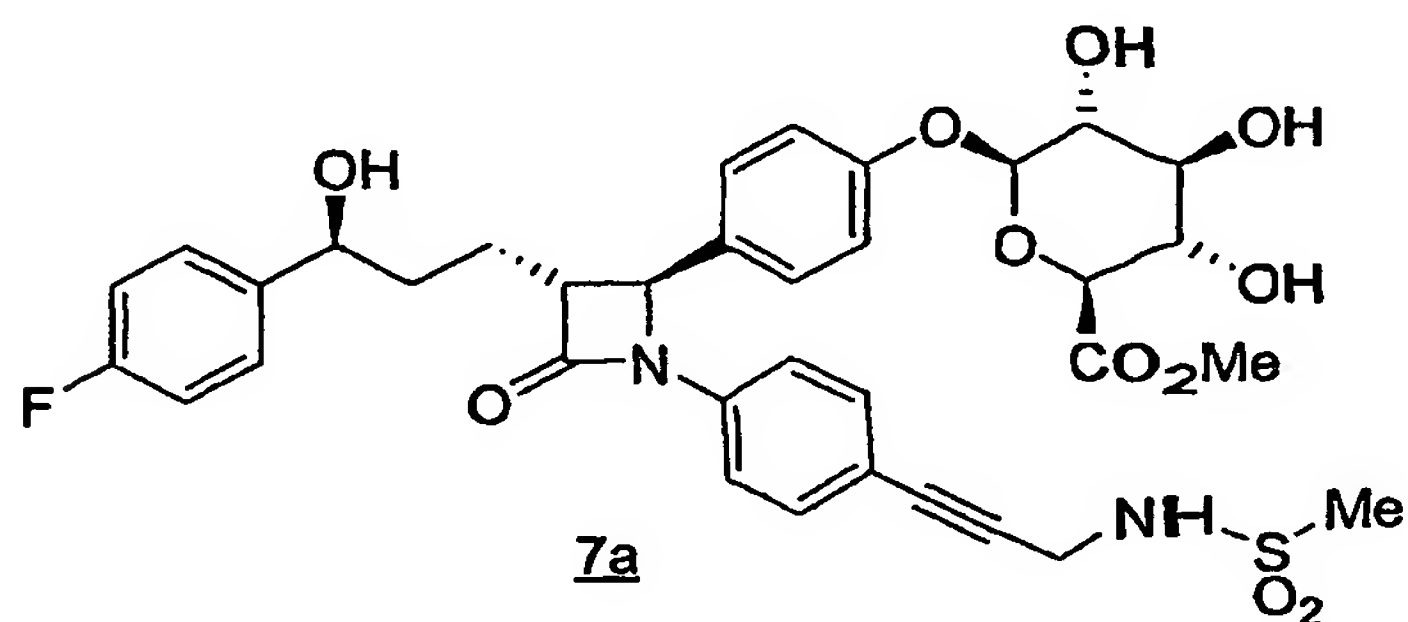


Triethylamine (7 equivalents) is added to a solution of (3*R*,4*S*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-4-(4-hydroxyphenyl)-1-(4-iodophenyl)azetidin-2-one (*i*-6) (1.00 equivalent), *N*-prop-2-yn-1-ylmethanesulfonamide (*i*-5) (1.50 equivalents), tetrakis(triphenylphosphine) palladium(0) (0.15 equivalents) and copper(I) iodide (0.30 equivalents) in DMF (0.1 M concentration with respect to final product) under a nitrogen atmosphere and the resulting solution aged at room temperature. After completion of reaction, the volatiles are evaporated in vacuo and the crude residue can be purified by flash chromatography on silica gel to afford the title compound.

EXAMPLE 44

Step A: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-{3-[(methylsulfonyl)amino]prop-1-yn-1-yl}phenyl)-4-oxoazetidin-2-yl]phenyl methyl β-D-glucopyranosiduronate (Compound 7a)

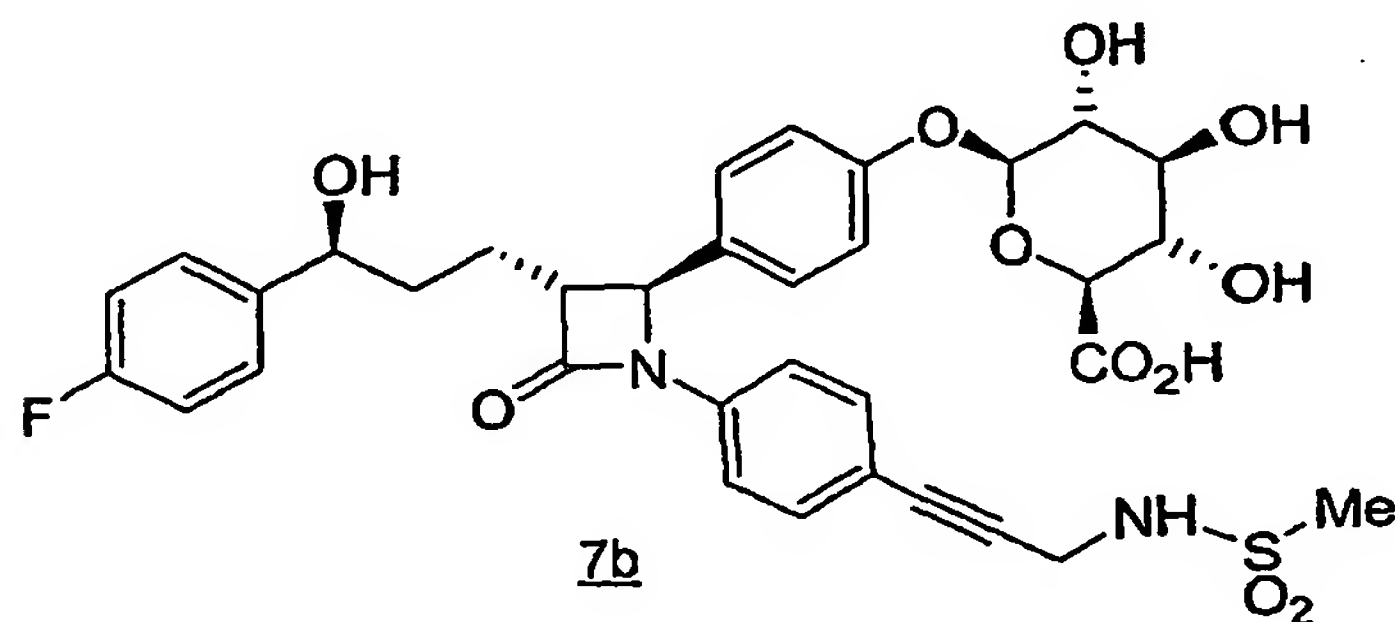
5



Triethylamine (0.07 mL, 0.502 mmol) was added to a stirred solution of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-iodophenyl)-4-oxoazetidin-2-yl]phenyl methyl β-D-glucopyranosiduronate (**i-7**) (0.050 g, 0.071 mmol), *N*-prop-2-yn-1-ylmethanesulfonamide (**i-5**) (0.014 g, 0.105 mmol), tetrakis(triphenylphosphine) palladium(0) (0.012 g, 0.010 mmol) and copper iodide (0.005 g, 0.026 mmol) in DMF (0.5 mL) under a nitrogen atmosphere and the resulting solution aged at room temperature for 18 h. The volatiles were evaporated in vacuo and the crude residue purified by flash chromatography on silica gel (gradient elution; 0-25% methanol/methylene chloride as eluent) to afford the title compound; *m/z* (ES) 713 (MH^+), 505.

15

Step B: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-{3-[(methylsulfonyl)amino]prop-1-yn-1-yl}phenyl)-4-oxoazetidin-2-yl]phenyl β-D-glucopyranosiduronic acid (Compound 7b, also referred to herein as compound 2)



5

A solution of compound 7a in methanol/water/triethylamine (1:7:2; 1 mL) was stirred at room temperature for approximately 1.5 h. The volatiles were evaporated *in vacuo* and the crude residue purified by preparative reversed phase high performance liquid chromatography on YMC Pack Pro C18 phase (gradient elution; 10-65% acetonitrile/water as eluent, 0.1% TFA modifier) to give the title compound (7b); *m/z* (ES) 699 (MH^+), 505; HRMS (ES) *m/z* calcd for $C_{34}H_{36}FN_2O_{11}S$ (MH^+) 699.2024, found 699.2016.

10

EXAMPLE 45 (COMPOUNDS 6B TO 6G AND 7C TO 7N)

The following compounds of Formula IIa have been prepared (as indicated by MS data provided) or can be prepared using the general synthetic procedures described in Example 43 (shown in Table 13) or Example 44 (shown in Table 14).

15

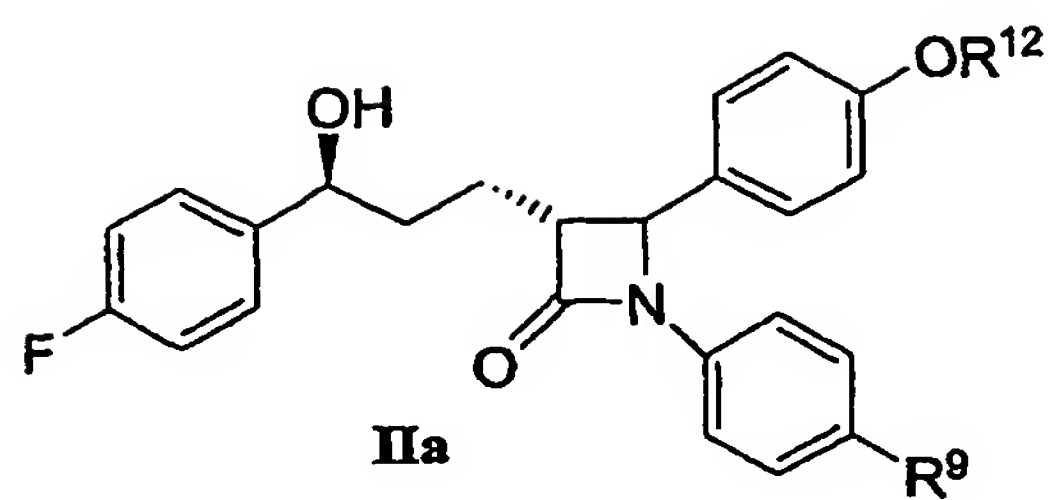


TABLE 13

Compound	R ¹²	R ⁹
6b)	H	
6c)	H	
6d)	H	
6e)	H	
6f)	H	
6g)	H	

TABLE 14

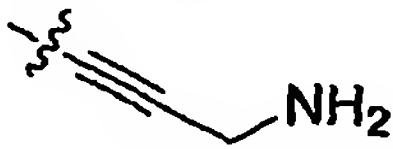

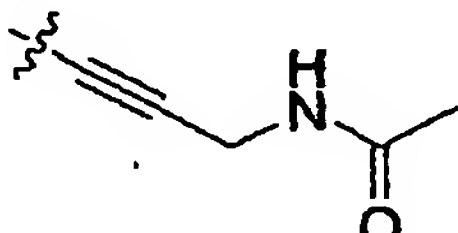
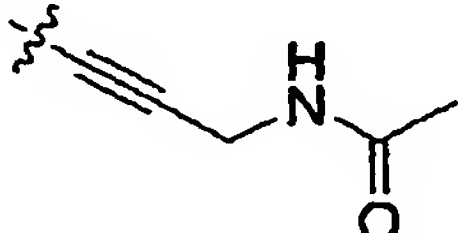
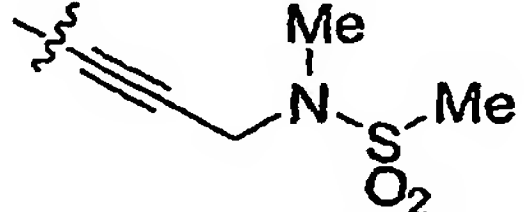
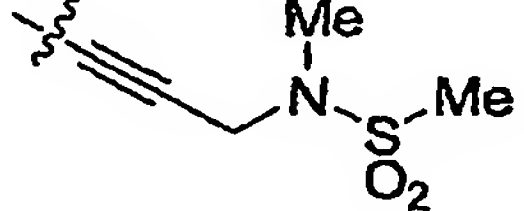
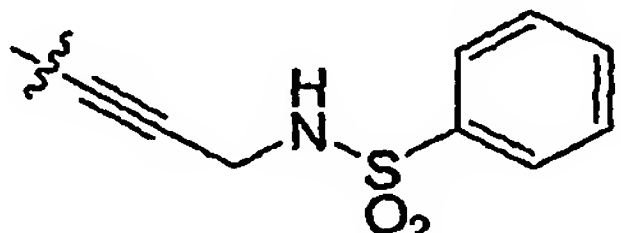
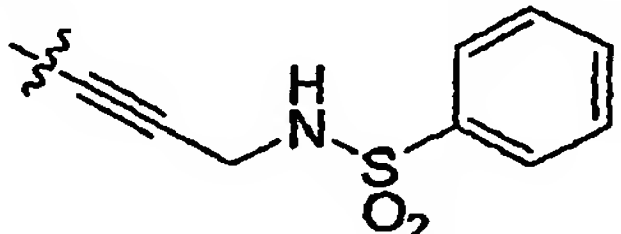
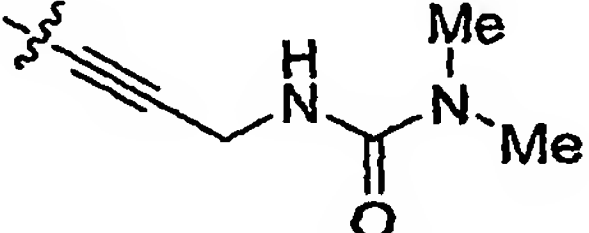
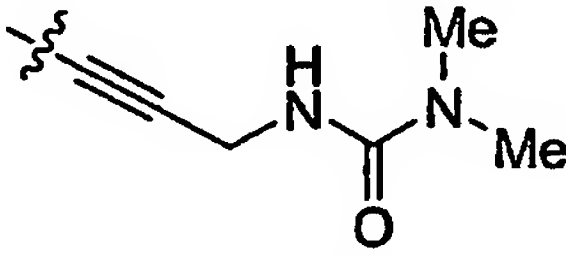
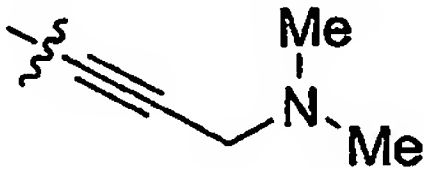
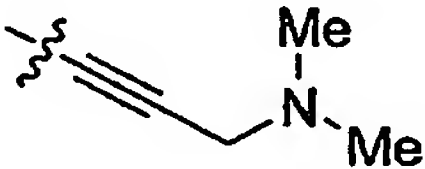
Compound	R12	R9	m/z (ES)	HRMS m/z (ES)
7c)	methyl ester glucuronide		658 (MNa ⁺)	
7d)	glucuronide		621 (MH ⁺)	(MH ⁺) Calcd 621.2249 Found 621.2223
7e)	methyl ester glucuronide		677 (MH ⁺)	
7f)	glucuronide		663 (MH ⁺)	(MH ⁺) Calcd 663.2354 Found 663.2331
7g)	methyl ester glucuronide		749 (MNa ⁺)	
7h)	glucuronide		735 (MNa ⁺)	(MH ⁺) Calcd 713.2180 Found 713.2170
7i)	methyl ester glucuronide		797 (MNa ⁺)	
7j)	glucuronide		783 (MNa ⁺)	(MH ⁺) Calcd 761.2180 Found 761.2193
7k)	methyl ester glucuronide		706 (MH ⁺)	

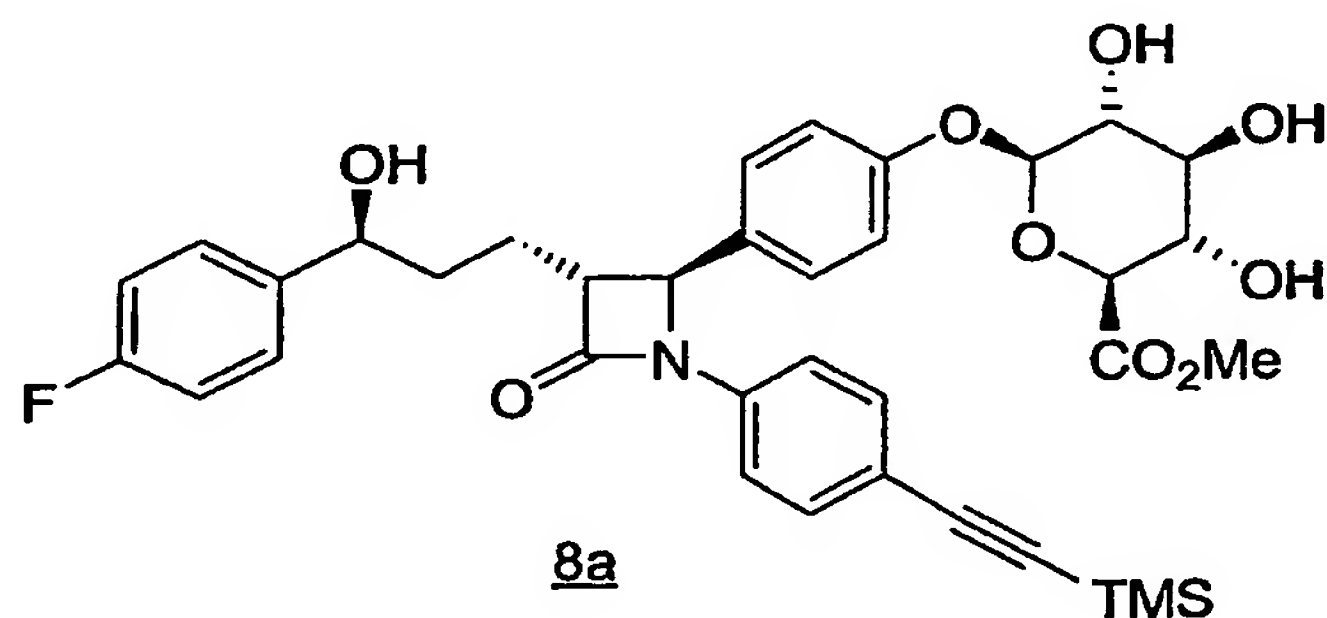
TABLE 14

Compound	R12	R9	m/z (ES)	HRMS m/z (ES)
7l)	glucuronide		692 (MH ⁺)	(MH ⁺) Calcd 692.2620 Found 692.2618
7m)	methyl ester glucuronide		663 (MH ⁺)	
7n)	glucuronide		649 (MH ⁺)	(MH ⁺) Calcd 649.2562 Found 649.2532

EXAMPLE 46

Step A: Preparation of 4-((2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-4-oxo-1-{4-[(trimethylsilyl)ethynyl]phenyl}azetidin-2-yl)phenyl methyl β-D-glucopyranosiduronate (8a)

5

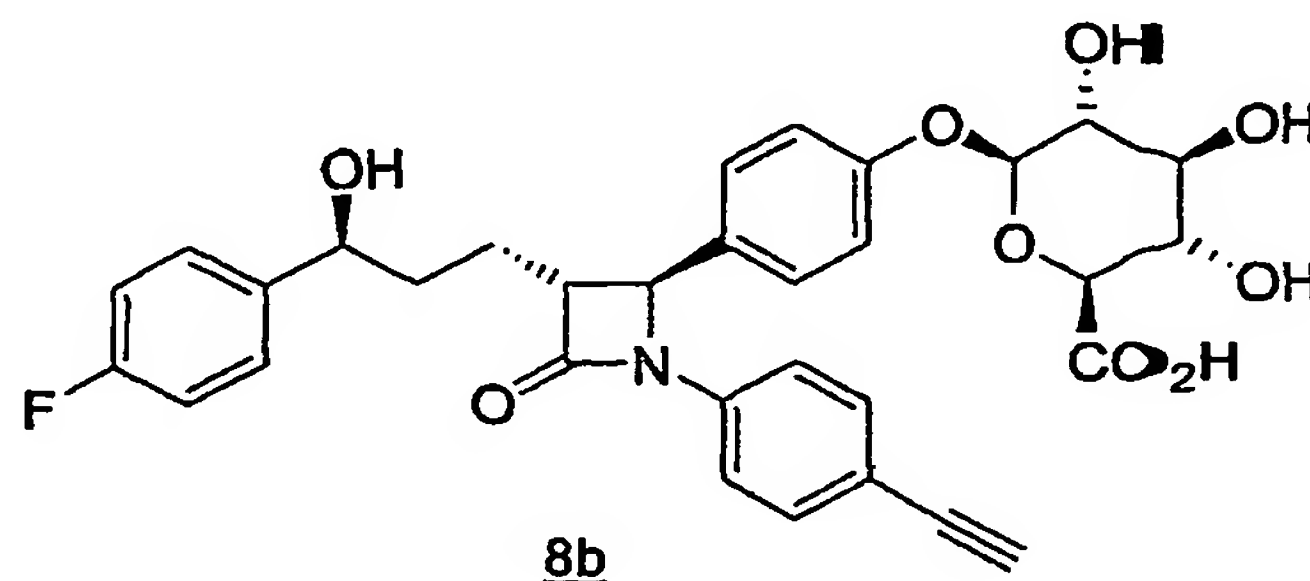


Triethylamine (69.0 μL, 0.495 mmol) was added to a stirred solution of i-7 (50.0 mg, 0.071 mmol), trimethylsilylacetylene (12.0 μL, 0.085 mmol), tetrakis(triphenylphosphine) palladium(0) (13.0 mg, 0.011 mmol) and copper iodide (5.10 mg, 0.028 mmol) in DMF (0.5 mL) under a nitrogen atmosphere and the resulting solution aged at room temperature for 18 h. The volatiles were evaporated in vacuo and the crude residue purified by flash chromatography on silica gel

10

(gradient elution; 0-25% methanol/methylene chloride as eluent) to afford the title compound (**8a**); m/z (ES) 660 (M-OH)⁺, 470.

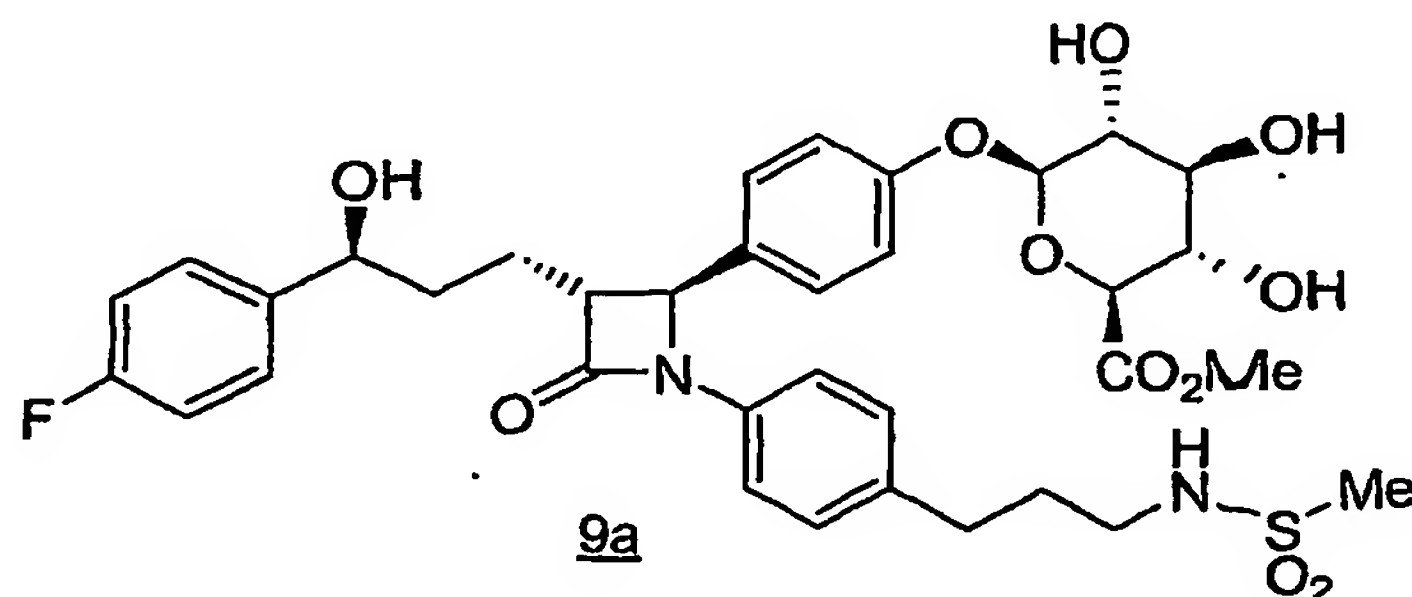
Step B: Preparation of 4-[(2*S*,3*R*)-1-(4-ethynylphenyl)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-4-oxoazetidin-2-yl]phenyl β-D-glucopyranosiduronic acid (**8b**)



A solution of **8a** in methanol/water/triethylamine (0.25 mL:1.10 mL:0.40 mL) was stirred at room temperature for approximately 6 h. The volatiles were evaporated *in vacuo* and the crude residue purified by preparative reversed phase high performance liquid chromatography on YMC Pack Pro C18 phase (gradient elution; 10-65% acetonitrile/water as eluent, 0.1% TFA modifier) to give the title compound (**8b**); m/z (ES) 574 (M-OH)⁺, 398; HRMS (ES) m/z calc'd for C₃₂H₃₁FNO₉ (MH⁺) 592.1983, found 592.1985.

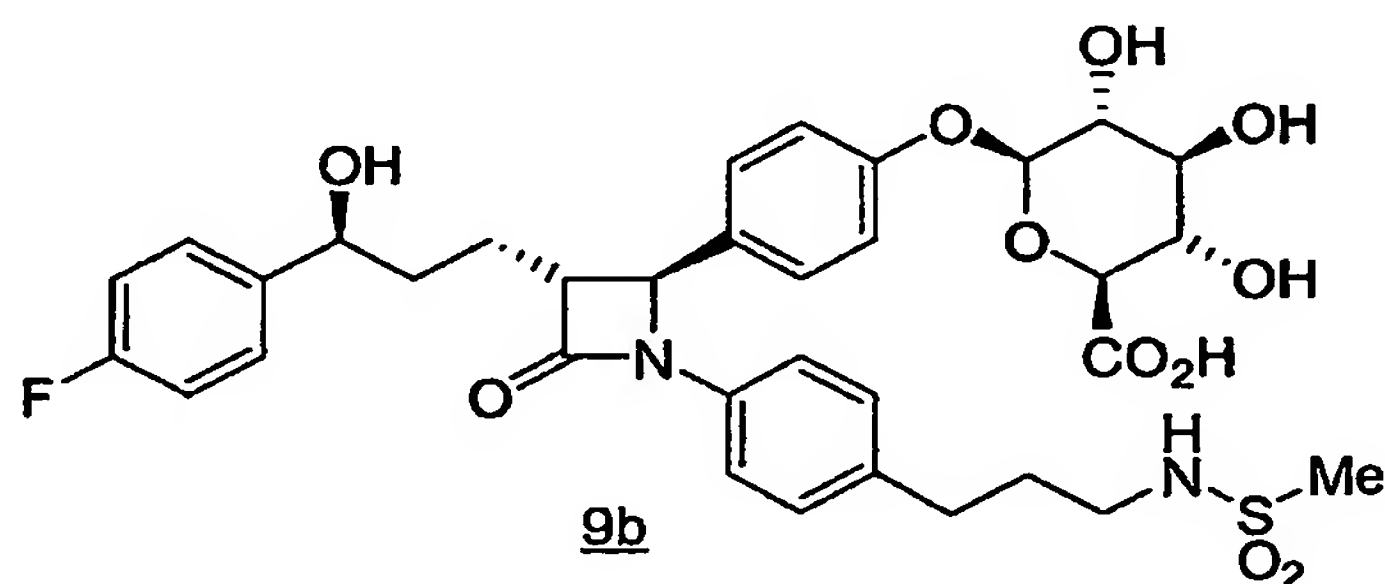
EXAMPLE 47

Step A: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-{3-[(methylsulfonyl)amino]propyl}phenyl)-4-oxoazetidin-2-yl]phenyl methyl β-D-glucopyranosiduronate (**9a**)



A mixture of **7a** (40.0 mg, 0.056 mmol) and palladium (~8 mg of 10 wt. % (dry basis) on activated carbon) in methanol (2 mL) was hydrogenated at atmospheric pressure for approximately 1 h. The reaction mixture was filtered through a short plug of celite, eluting copiously with methanol, and the filtrate
 5 evaporated *in vacuo* to afford the title compound (**9a**); *m/z* (ES) 509 (M-sugar-OH)⁺.

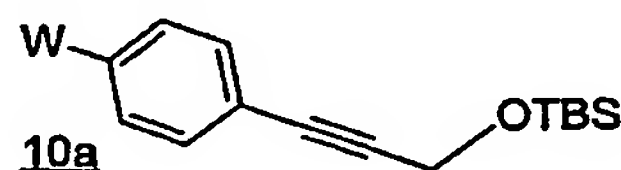
Step B: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-1-(4-{3-[(methylsulfonyl)amino]propyl}phenyl)-4-oxoazetidin-2-yl]phenyl β-D-glucopyranosiduronic acid (**9b**)



10 A solution of **9a** in methanol/water/triethylamine (1:7:2, 1 mL) was stirred at room temperature for approximately 1 h. The volatiles were evaporated *in vacuo* and the crude residue purified by preparative reversed phase high performance liquid chromatography on YMC Pack Pro C18 phase (gradient elution; 10-65% acetonitrile/water as eluent, 0.1% TFA modifier) to give the title compound (**9b**); *m/z*
 15 (ES) 735 (M+Na)⁺, 685 (M-OH)⁺, 509 (M-sugar-OH)⁺; HRMS (ES) *m/z* calc'd for C₃₄H₃₉FN₂O₁₁S (MH⁺) 703.2337, found 703.2337.

EXAMPLE 48

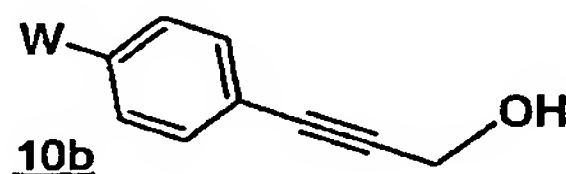
Step A: Preparation of 4-{(2*S*,3*R*)-3-[(3*S*)-3-acetoxy)-3-(4-fluorophenyl)propyl]-1-[4-(3-{[*tert*-butyl(dimethylsilyl)oxy]prop-1-yn-1-yl}phenyl]-4-oxoazetidin-2-yl}phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (**10a**)



Triethylamine (170 μL, 1.25 mmol) was added to a solution of **i-8** (156 mg, 0.178 mmol), *tert*-butyldimethyl(2-propynyloxy)silane (43.0 μL, 0.214 mmol),

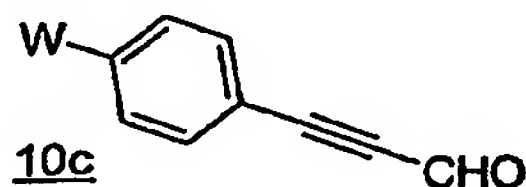
dichlorobis(triphenylphosphine) palladium(II) (12.0 mg, 0.018 mmol) and copper iodide (7.00 mg, 0.036 mmol) in DMF (1.3 mL) under a nitrogen atmosphere and the resulting solution aged at room temperature for approximately 20 h. The reaction mixture was poured into saturated aqueous sodium bicarbonate and extracted twice with diethyl ether. The combined organic extracts were washed with water, brine, dried (MgSO₄), filtered and the filtrate concentrated *in vacuo*. Purification of the crude residue by flash chromatography on silica gel (gradient elution; 15-40% ethyl acetate/hexanes as eluent) afforded the title compound 10a.

Step B: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-1-[4-(3-hydroxyprop-1-yn-1-yl)phenyl]-4-oxoazetidin-2-yl}phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (10b)



Tetrabutylammonium fluoride hydrate (39.0 mg, 0.148 mmol) was added to 10a (136 mg, 0.148 mmol) in tetrahydrofuran (1.5mL), and the resulting solution aged at room temperature for 1 h. The reaction mixture was poured into saturated aqueous ammonium chloride and extracted twice with ether. The combined organic extracts were washed with saturated sodium bicarbonate, brine, dried (MgSO₄), filtered and concentrated *in vacuo*. Purification of the crude residue by flash chromatography on silica gel (50% ethyl acetate/hexanes) afforded the title compound 10b.

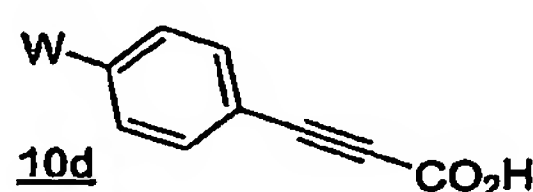
Step C: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-4-oxo-1-[4-(3-oxoprop-1-yn-1-yl)phenyl]azetidin-2-yl}phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (10c)



Dess-Martin periodinane (33.0 mg, 0.077mmol) was added to a solution of 10b (62.0 mg, 0.077 mmol) and pyridine (31.0 μL, 0.386 mmol) in dichloromethane (1 mL) at room temperature. After 30 min, the reaction mixture was

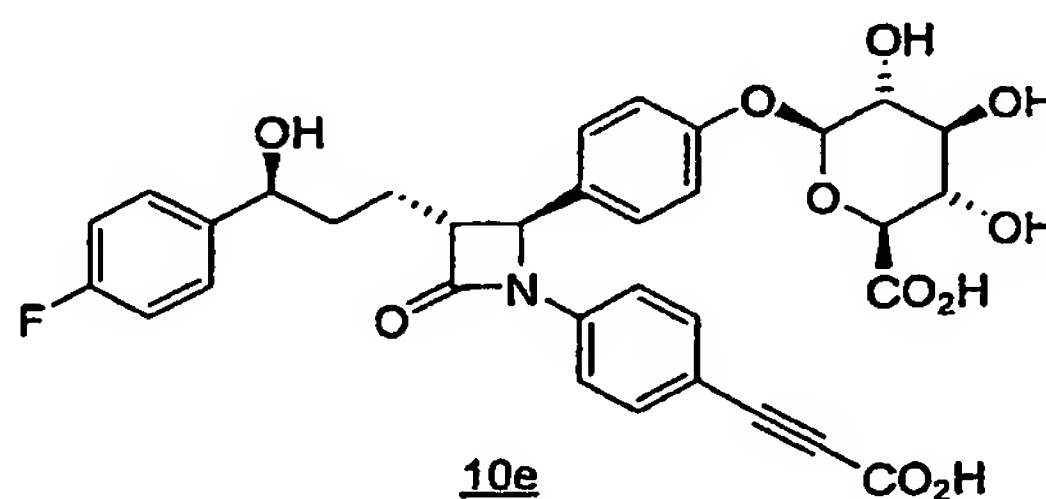
poured into saturated aqueous sodium bicarbonate, and extracted twice with ethyl acetate. The combined organic extracts were washed with water, brine, dried (MgSO₄), filtered and concentrated *in vacuo*. Purification of the crude residue by flash chromatography on silica gel (gradient elution; 20-40% ethyl acetate/hexanes) afforded the title compound 10c.

Step D: Preparation of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-1-[4-(carboxyethynyl)phenyl]-4-oxoazetidin-2-yl]phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (10d)



An aqueous solution (0.1 mL) of sodium dihydrogenphosphate (9.00 mg, 0.065 mmol) and sodium chlorite (5.00 mg, 0.055 mmol) was added to a solution of 10c (37.0 mg, 0.046 mmol) in *tert*-butyl alcohol (0.4 mL), dioxane (0.2 mL) and isobutylene (~0.1 mL) at room temperature. After 1.5 h, the reaction mixture was concentrated *in vacuo* and the crude residue triturated repeatedly with ethyl acetate. The organic washings were dried (Na₂SO₄), filtered and concentrated *in vacuo* to afford the title compound 10d.

Step E: Preparation of 4-[(2*S*,3*R*)-1-[4-(carboxyethynyl)phenyl]-3-[(3*S*)-3-(4-fluorophenyl)-3-hydroxypropyl]-4-oxoazetidin-2-yl]phenyl β-D-glucopyranosiduronic acid (10e)

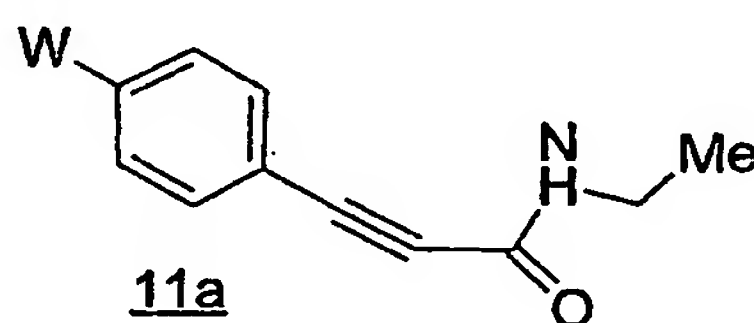


A solution of 4-[(2*S*,3*R*)-3-[(3*S*)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-1-[4-(carboxyethynyl)phenyl]-4-oxoazetidin-2-yl]phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (10d) and sodium cyanide (~1mg, 0.020 mmol) in methanol (3mL) was heated to 45°C. After 22 h, the reaction mixture was concentrated under reduced pressure and dissolved in methanol/water/triethylamine (1:7:2, 1 mL). After stirring at room temperature for

approximately 1 h, the volatiles were evaporated *in vacuo* and the crude residue purified by preparative reversed phase high performance liquid chromatography on YMC Pack Pro C18 phase (gradient elution; 10-60% acetonitrile/water as eluent, 0.1% TFA modifier) to give the title compound (10e), *m/z* (ES) 442.0 (M-sugar-OH)⁺, 618.0 (M-OH)⁺; HRMS (ES) *m/z* calcd. for C₃₃H₃₁FNO₁₁ (MH⁺) 636.1881, found 636.1889

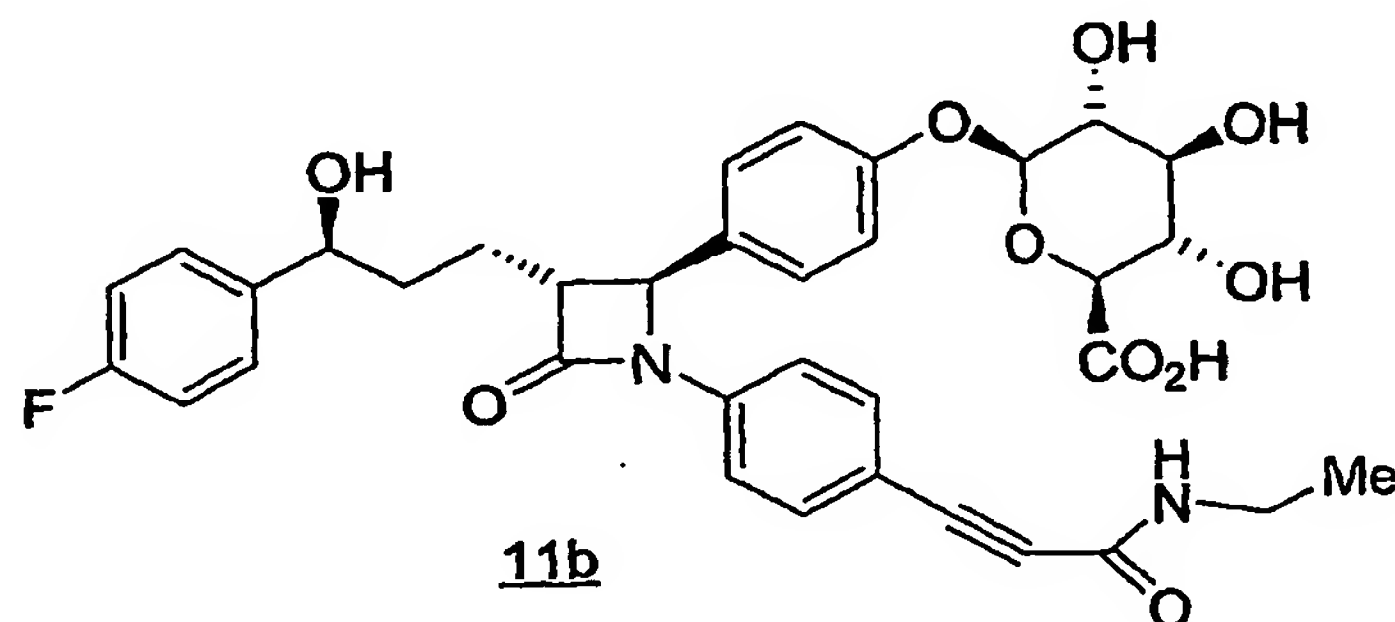
EXAMPLE 49

Step A: Preparation of 4-((2*S*,3*R*)-3-[(3*S*)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-1-{4-(3-(ethylamino)-3-oxoprop-1-yn-1-yl)phenyl}-4-oxoazetidin-2-yl)phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (11a)



A 1M solution of ethylamine hydrochloride and diisopropylethylamine in DMF (40.0 μL, 0.40 mmol) was added to 4-((2*S*,3*R*)-3-[(3*S*)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-1-{4-(carboxyethynyl)phenyl}-4-oxoazetidin-2-yl}phenyl methyl 2,3,4-tri-*O*-acetyl-β-D-glucopyranosiduronate (10d) (27.0 mg, 0.033 mmol), 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (19.0 mg, 0.099 mmol) and 1-hydroxybenzotriazole (8.00 mg, 0.059 mmol) in DMF (0.25 mL). After 4.5 h, the reaction mixture was poured into ethyl acetate and washed successively with water and brine. The organic layer was dried, filtered and concentrated under reduced pressure. Purification of the crude residue by flash chromatography on silica gel (gradient elution; 50-60% ethyl acetate/hexanes) afforded the title compound 11a.

Step B: Preparation of 4-((2S,3R)-1-{4-[3-(ethylamino)-3-oxoprop-1-yn-1-yl]phenyl}-3-[(3S)-3-(4-fluorophenyl)-3-hydroxypropyl]-4-oxoazetidin-2-yl}phenyl β-D-glucopyranosiduronic acid (11b)



5 A solution of 4-((2S,3R)-3-[(3S)-3-(acetyloxy)-3-(4-fluorophenyl)propyl]-1-{4-(3-(ethylamino)-3-oxoprop-1-yn-1-yl]phenyl}-4-oxoazetidin-2-yl)phenyl methyl 2,3,4-tri-O-acetyl-β-D-glucopyranosiduronate (11a) (22.0 mg, 0.026 mmol) and sodium cyanide (~1mg, 0.020 mmol) in methanol (3mL) was heated to 45°C. After 18 h, the reaction mixture was concentrated under reduced
10 pressure and dissolved in methanol/water/triethylamine (1:3:1, 2.5 mL). After stirring at room temperature for approximately 1 h, the volatiles were evaporated *in vacuo*, and the crude residue purified by preparative reversed phase high performance liquid chromatography on YMC Pack Pro C18 phase (gradient elution; 10-60% acetonitrile/water as eluent, 0.1% TFA modifier) to give the title compound (11b) *m/z* (ES) 663.0 (M+H)⁺; HRMS (ES) *m/z* calcd. for C₃₅H₃₆FN₂O₁₀ (MH⁺) 663.2354, found 663.2341.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in
20 addition to those described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

Patents, patent applications, publications, product descriptions, Genbank Accession Numbers and protocols are cited throughout this application, the
25 disclosures of which are incorporated herein by reference in their entireties for all purposes.

CLAIMS

1. A method for identifying a ligand of NPC1L1 comprising:
contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound; and
determining whether said candidate compound binds to human NPC1L1;
wherein binding of said candidate compound to human NPC1L1 modulates binding of said detectably labeled substituted 2-azetidinone glucuronide to human NPC1L1, wherein the detectably labeled substituted 2-azetidinone glucuronide has a binding affinity K_D value for human NPC1L1 that is 200nM or lower, and wherein said modulation indicates that the candidate compound is a ligand that binds to human NPC1L1.
2. The method of claim 1, wherein the K_D value is 100nM or lower.
3. The method of claim 1, wherein the K_D value is 50nM or lower.
4. The method of claim 1, wherein the K_D value is 20nM or lower.
5. The method of claim 1, wherein the K_D value is 10nM or lower.
6. The method of claim 1, wherein the substituted 2-azetidinone-glucuronide is selected from the group consisting of a compound of Formula I and a compound of Formula II.
7. The method of claim 6, wherein the substituted 2-azetidinone-glucuronide comprises a detectable label from the group consisting of ^{35}S and ^{125}I .
8. The method of claim 7, wherein the detectable label is ^{35}S .
9. The method of claim 6, wherein the substituted 2-azetidinone-glucuronide is a compound of Formula II, wherein R^9 comprises an $-\text{SO}_2-$ group.
10. The method of claim 9, wherein the substituted 2-azetidinone-glucuronide of Formula II is labeled with ^{35}S .

11. A method for identifying a ligand of NPC1L1 comprising:
contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide of Formula II and a candidate compound; and
determining whether said candidate compound binds to human NPC1L1;
wherein binding of said candidate compound to human NPC1L1 modulates binding of said detectably labeled substituted 2-azetidinone glucuronide of Formula II to human NPC1L1, and wherein said modulation indicates that the candidate compound is a ligand that binds to human NPC1L1.
12. The method of claim 11, wherein R⁹ of the detectably labeled substituted 2-azetidinone glucuronide of Formula II comprises an -SO₂- group.
13. The method of claim 11, wherein the detectably labeled substituted 2-azetidinone glucuronide of Formula II is labeled with ³⁵S.
14. The method of claim 11, wherein the detectably labeled substituted 2-azetidinone glucuronide of Formula II has a binding affinity K_D value for human NPC1L1 that is 200nM or lower.
15. The method of claim 14, wherein the K_D value is 100nM or lower.
16. The method of claim 14, wherein the K_D value is 50nM or lower.
17. The method of claim 14, wherein the K_D value is 20nM or lower.
18. The method of claim 14, wherein the K_D value is 10nM or lower.
19. The method of claim 1 wherein the detectably labeled substituted 2-azetidinone glucuronide is labeled with ³⁵S.
20. The method of claim 1 wherein the detectably labeled substituted 2-azetidinone glucuronide is ³⁵S-labeled compound 2.

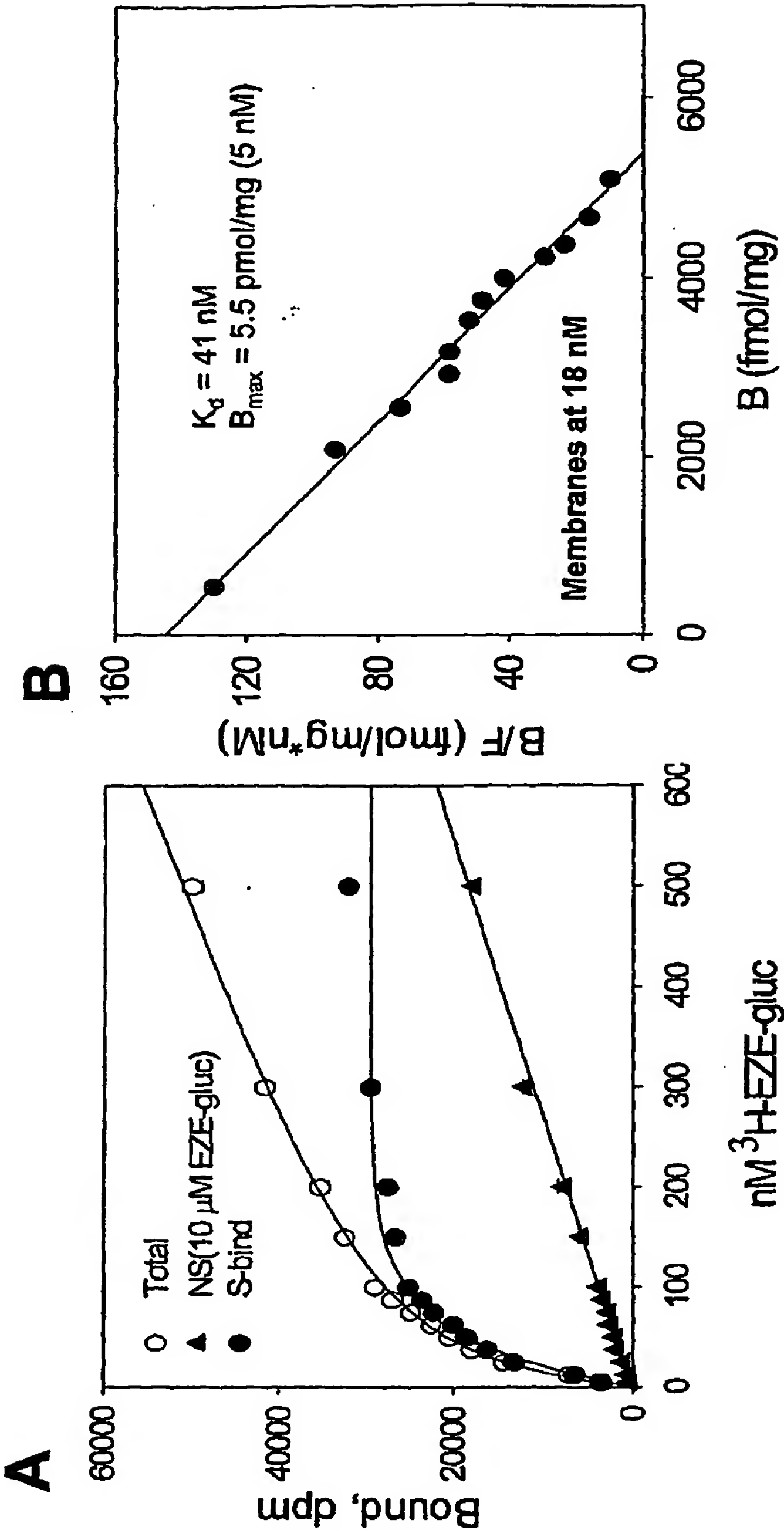


Figure 1. Equilibrium binding of EZE-glucuronide to rhesus BBMVs

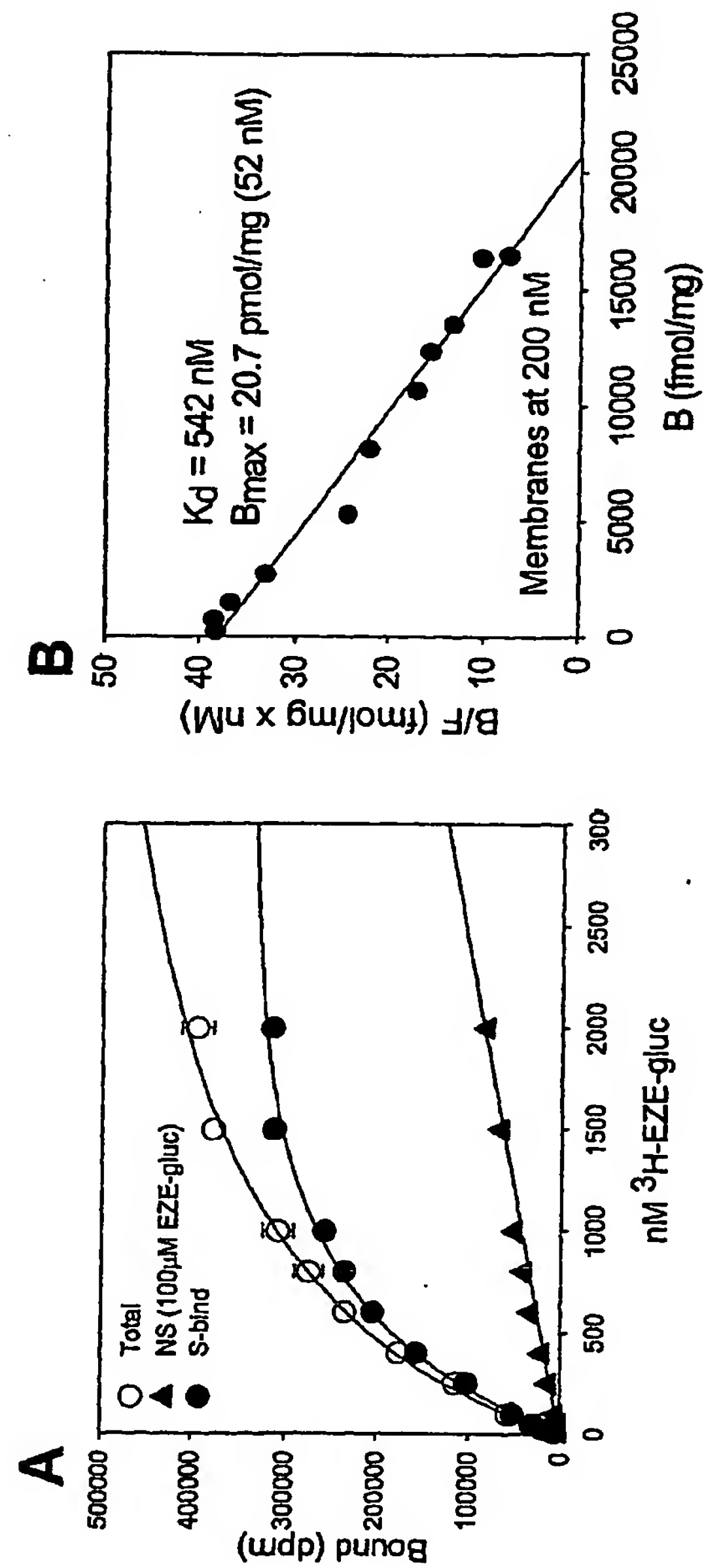


Figure 2. Equilibrium binding of EZE-glucuronide to rat BBMV.

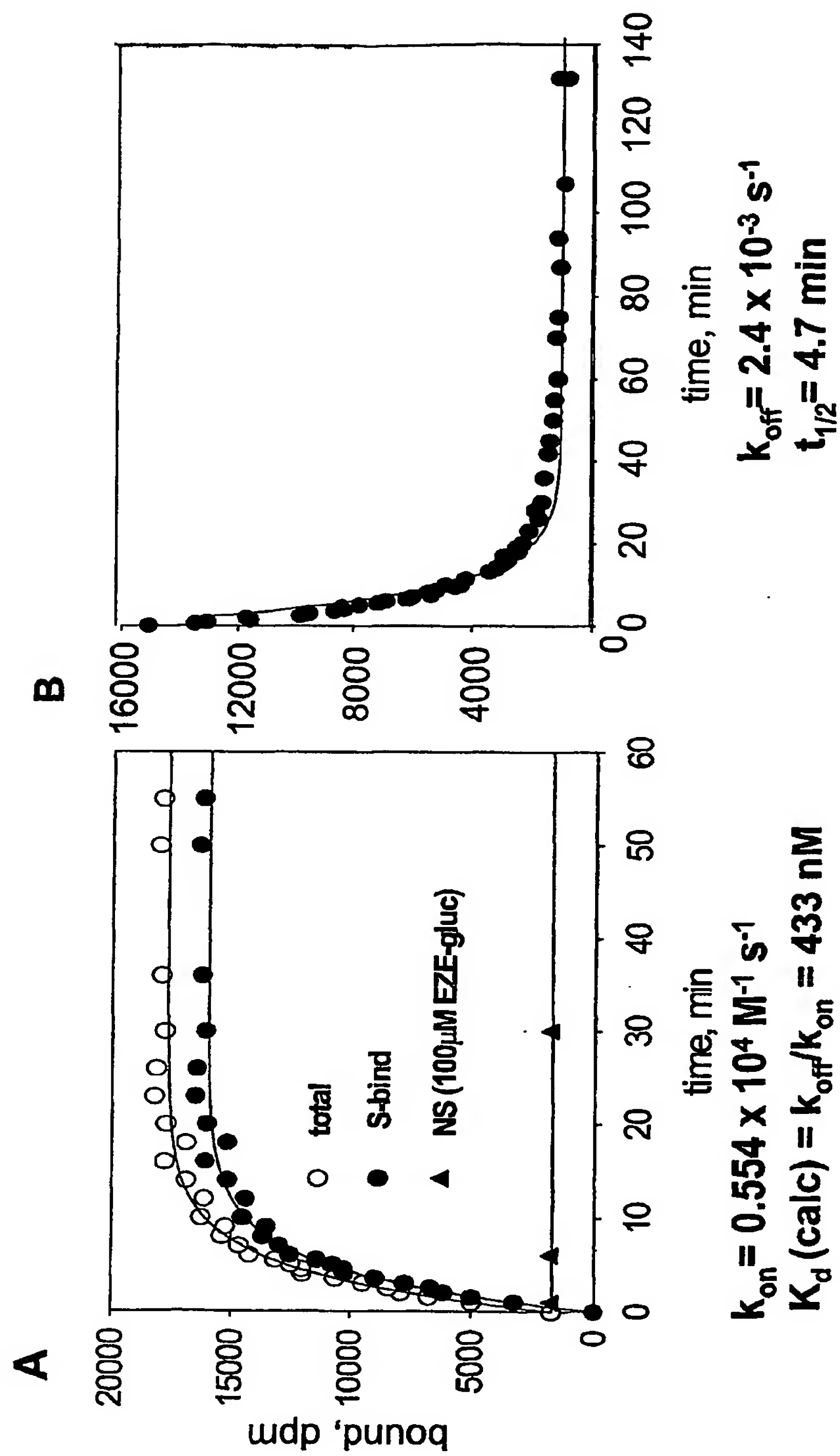


Figure 3. Association and dissociation kinetics of ³H-EZE-glucuronide in rat BBMV.

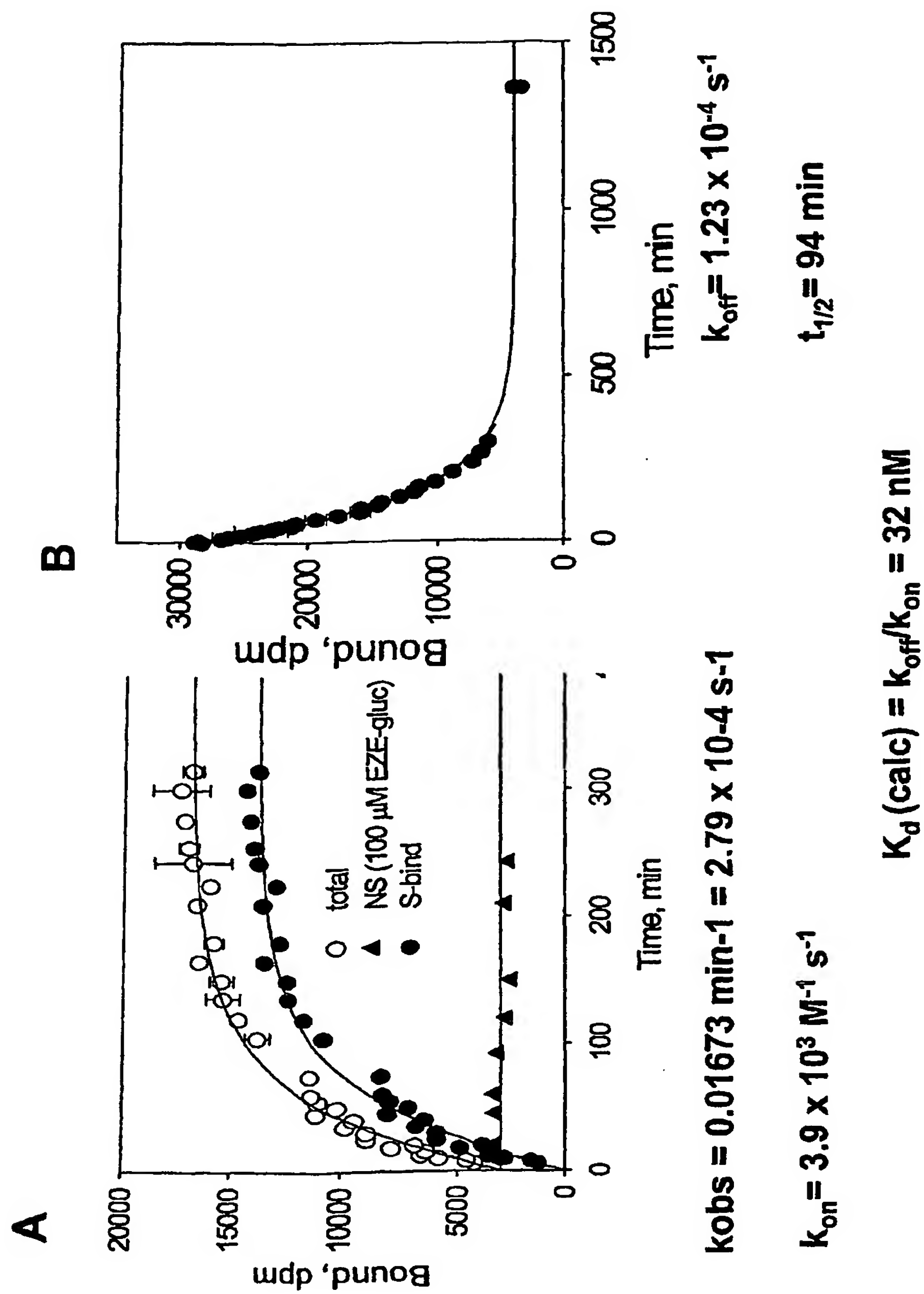


Figure 4. Association and Dissociation kinetics of ^3H -EZE-glucuronide in rhesus BBMV.

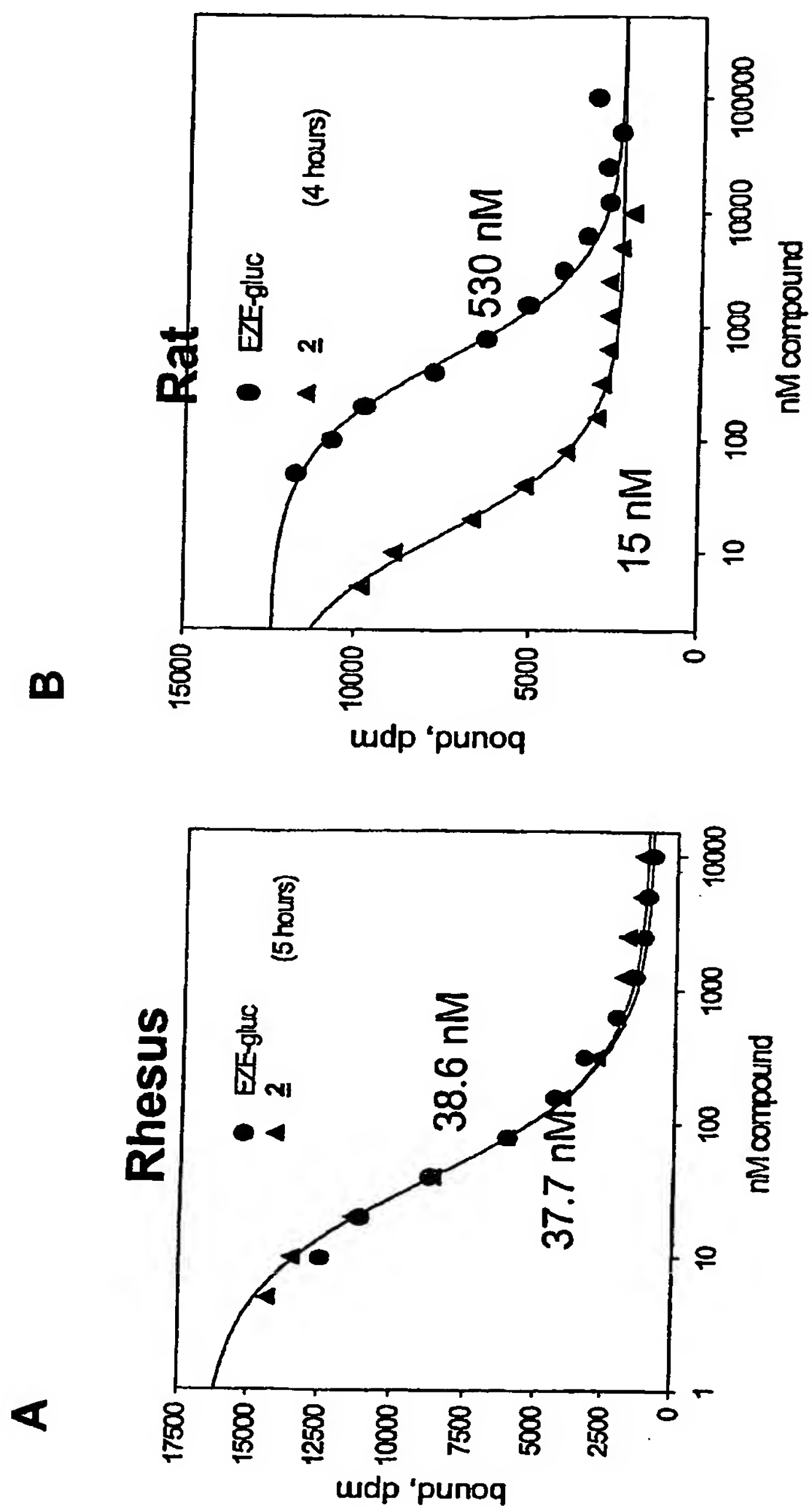


Figure 5. Displacement of 3H-EZE-glucuronide by EZE-glucuronide and compound 2 in rhesus and rat BBMV.

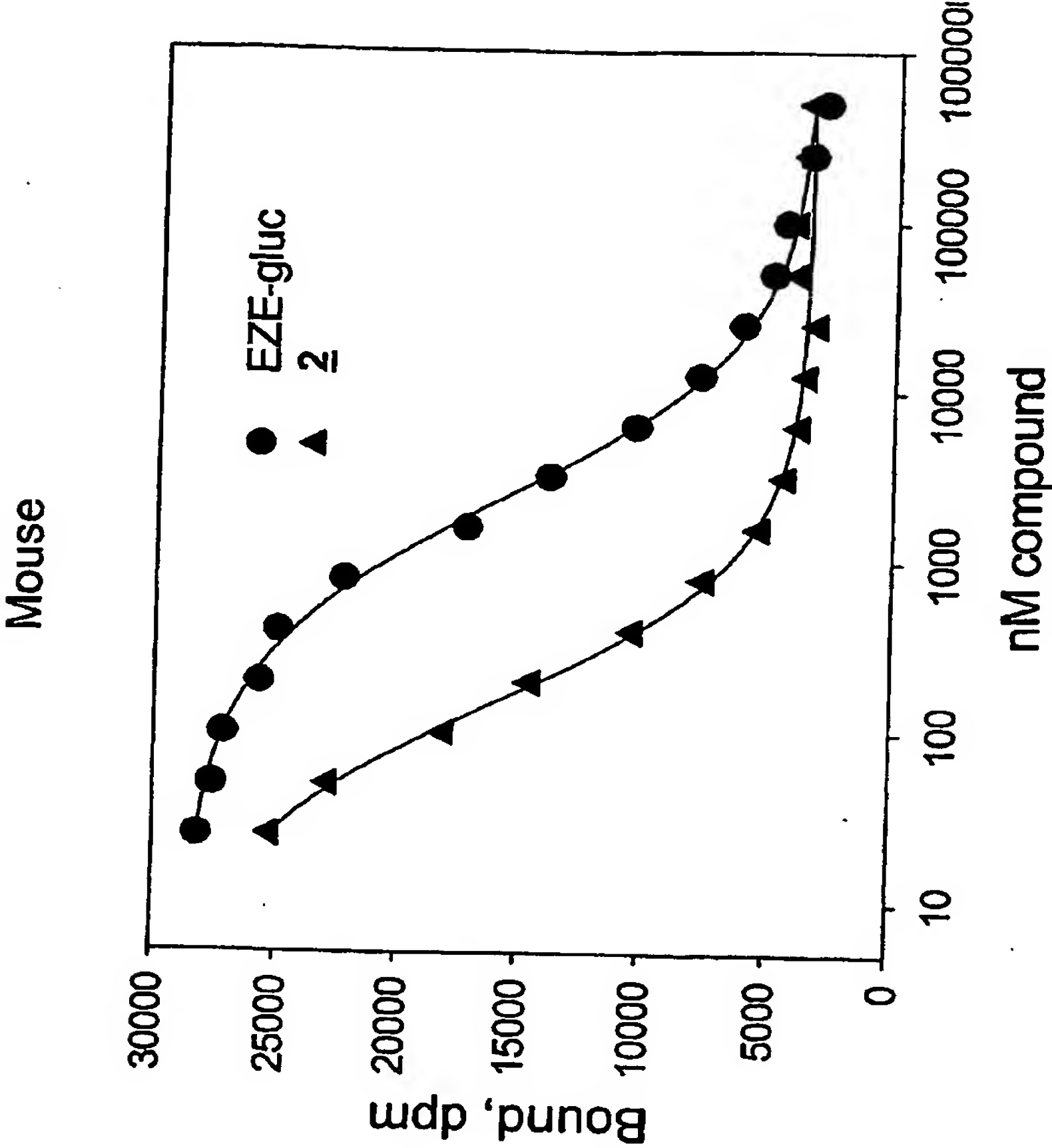


Figure 6. Displacement of ³⁵S-labeled compound 2 by EZE-glucuronide and compound 2 in mouse BBMV.

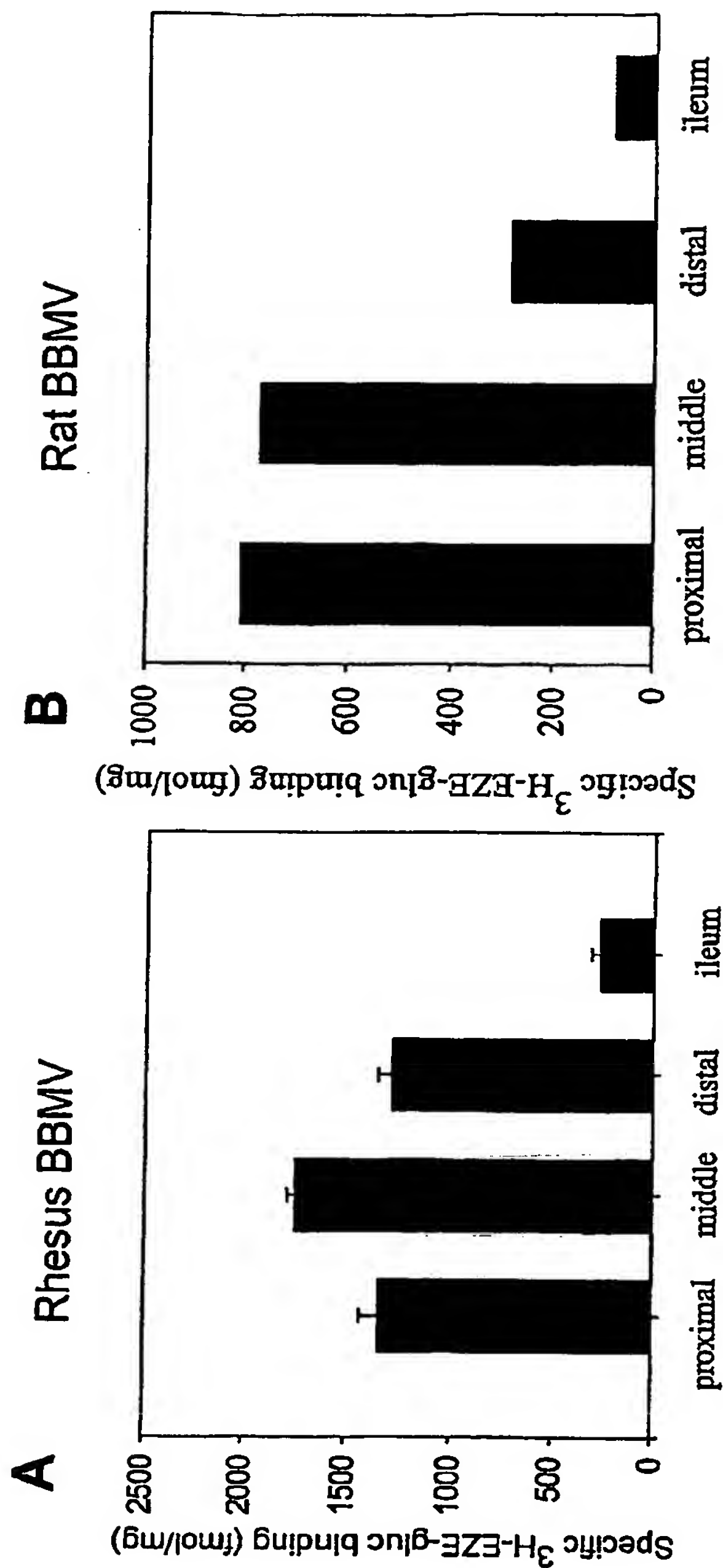


Figure 7. Intestinal distribution of ezetimibe binding sites.

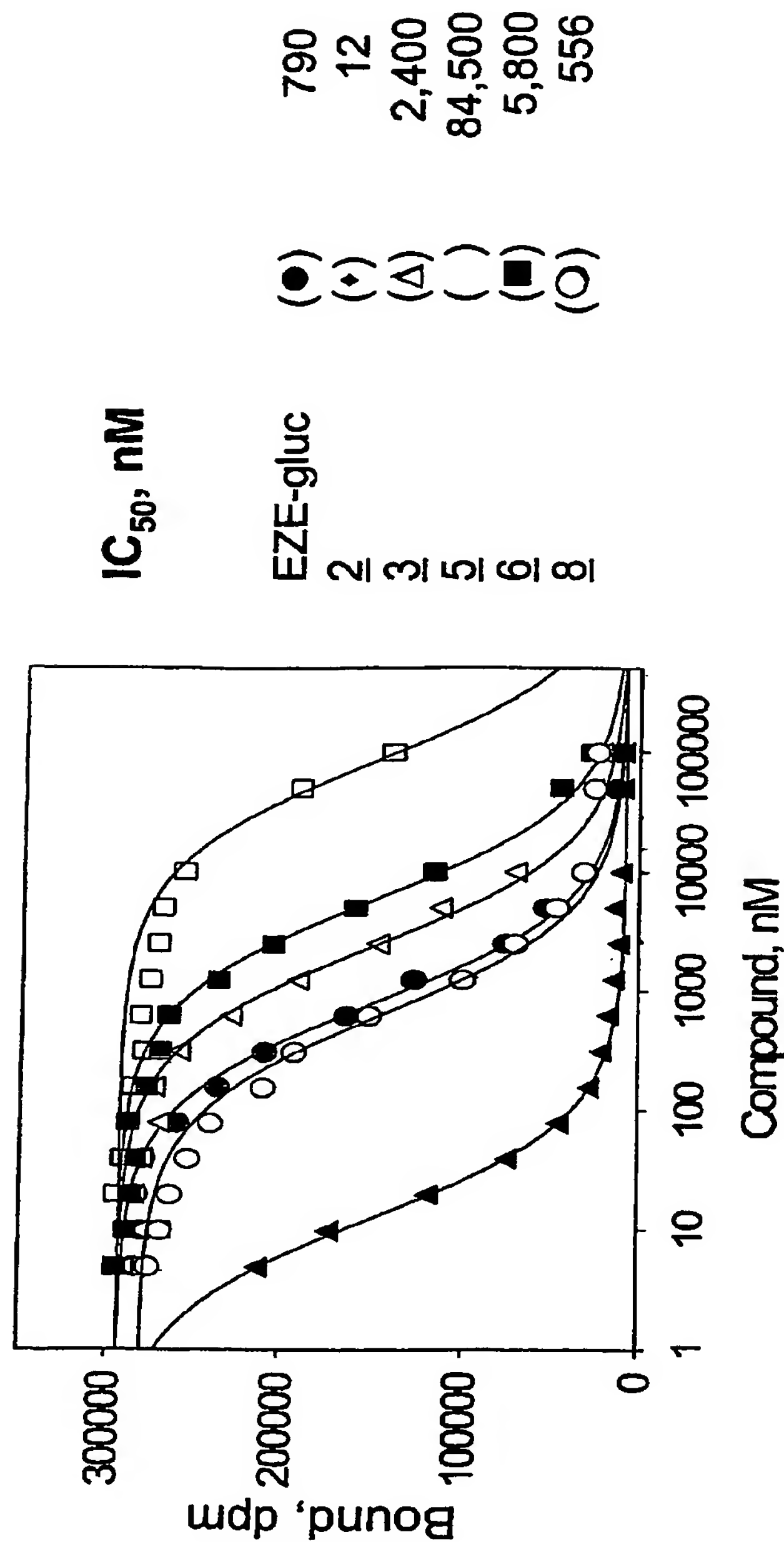


Figure 8. Displacement of ³⁵S-labeled compound 2 by EZE-glucuronide and analogs in transfected CHO cells expressing rat NPC1L1

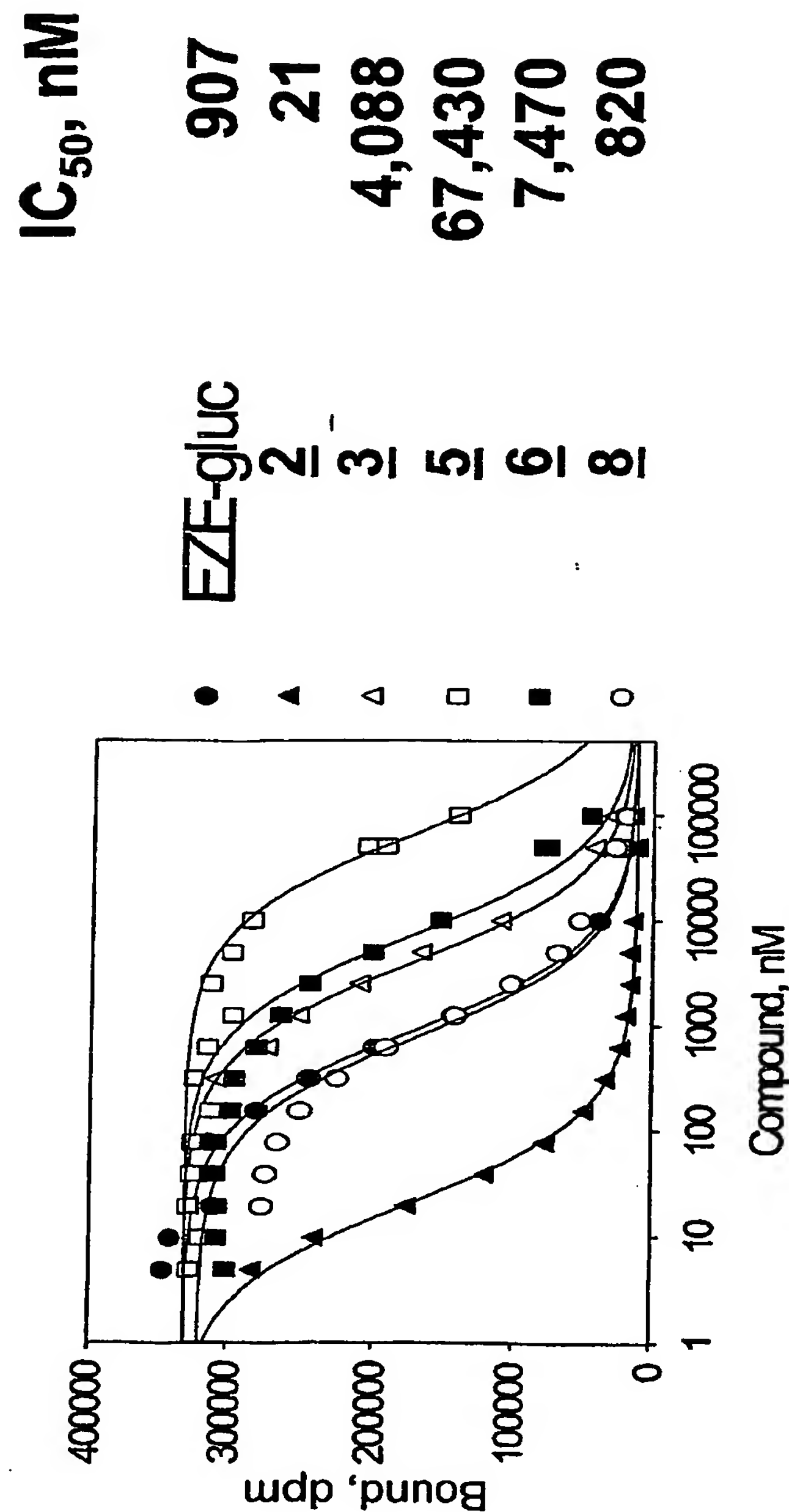


Figure 9. Displacement of ³⁵S-labeled compound 2 by EZE-glucuronide and analogs in transfected CHO cells expressing human NPC1L1

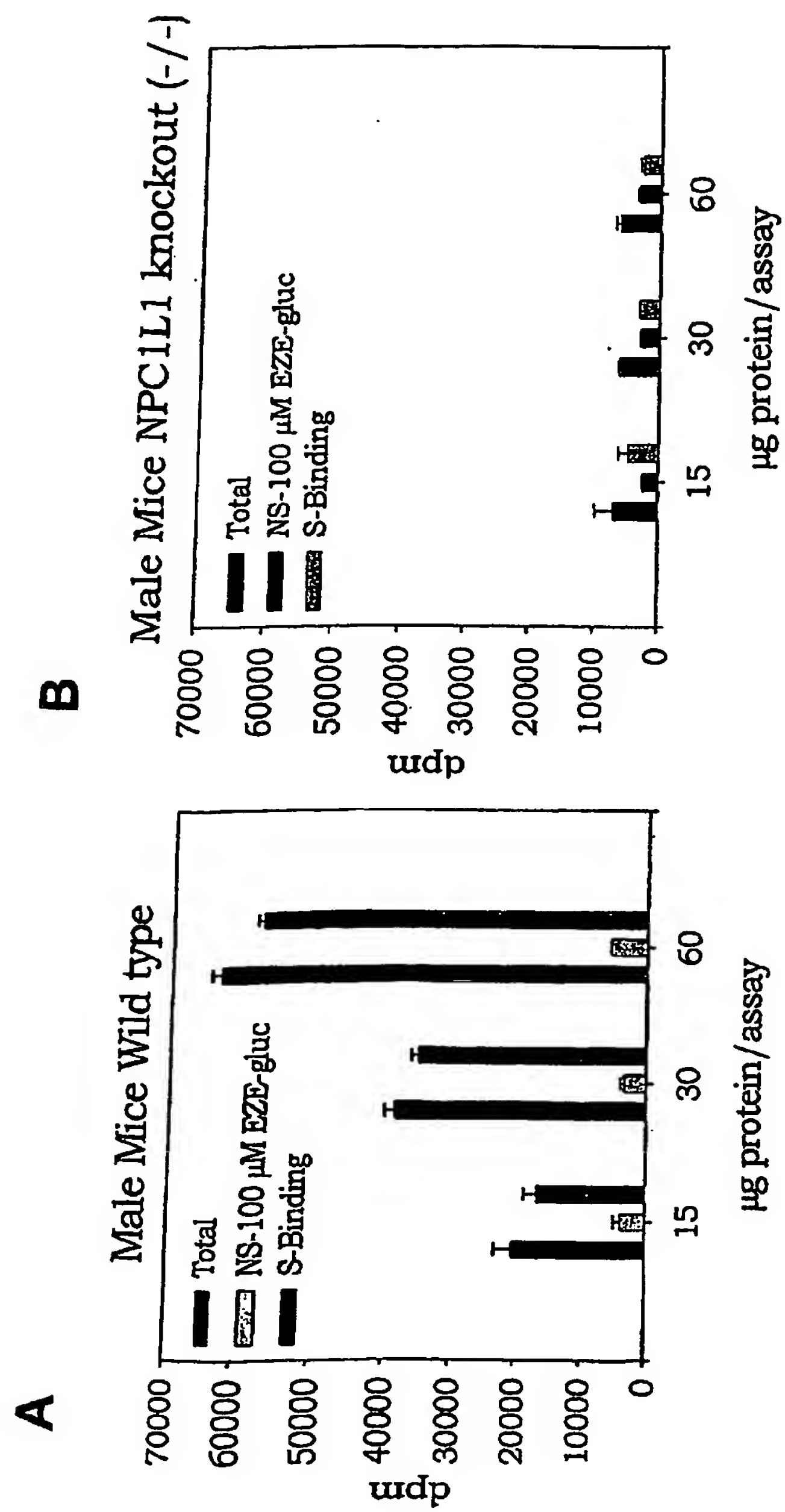


Figure 10. ^{35}S -labeled compound 2 binding with brush border membranes from intestinal mucosal scrapings of male wild type (A) and NPC1L1 knockout (-/-) mice (B).

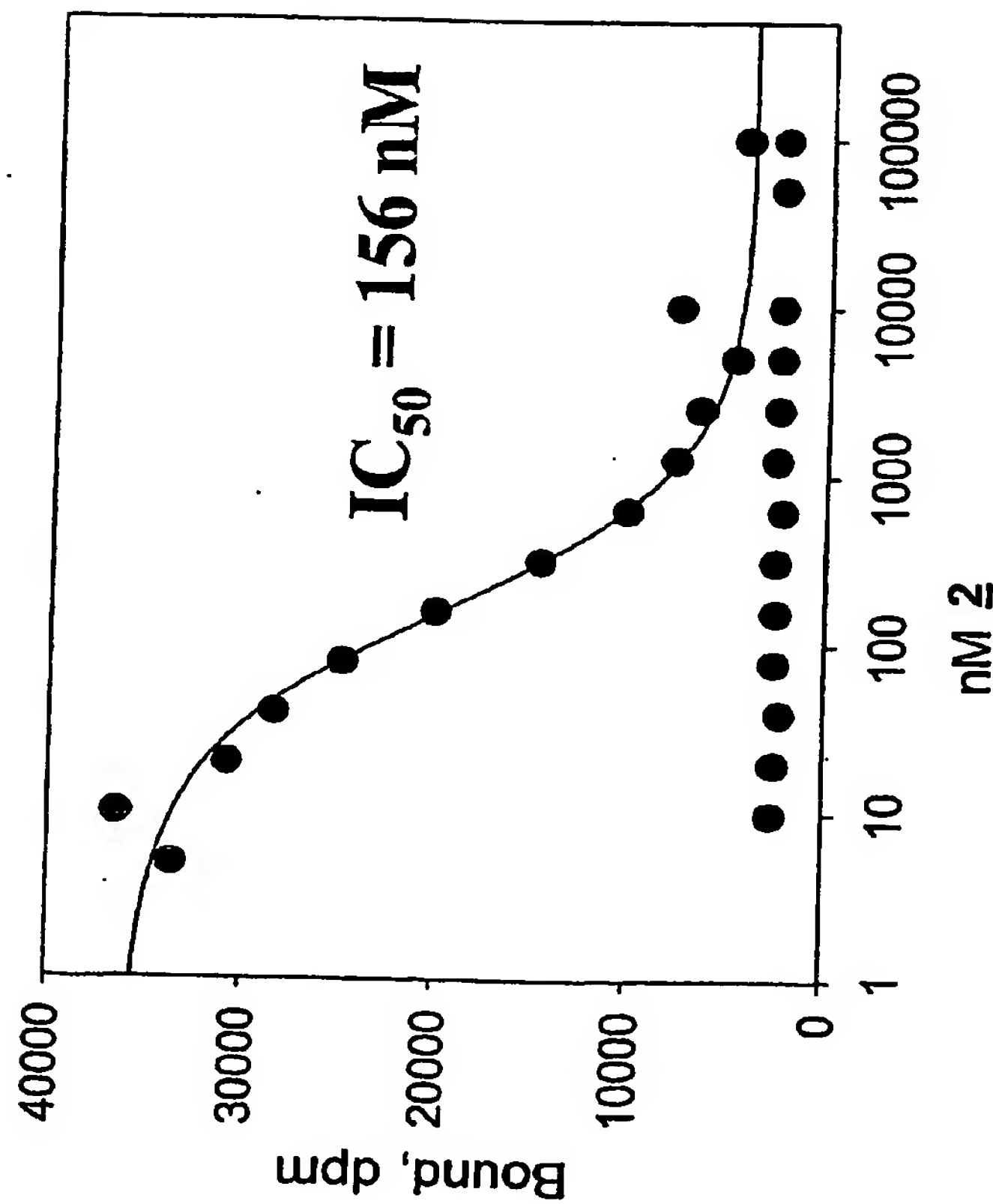


Figure 11. Displacement of ^{35}S -labeled compound 2 by compound 2 in mouse wild type and knockout mouse NPC1L1 (-/-) BBMV.

FIGURE 12

Competition

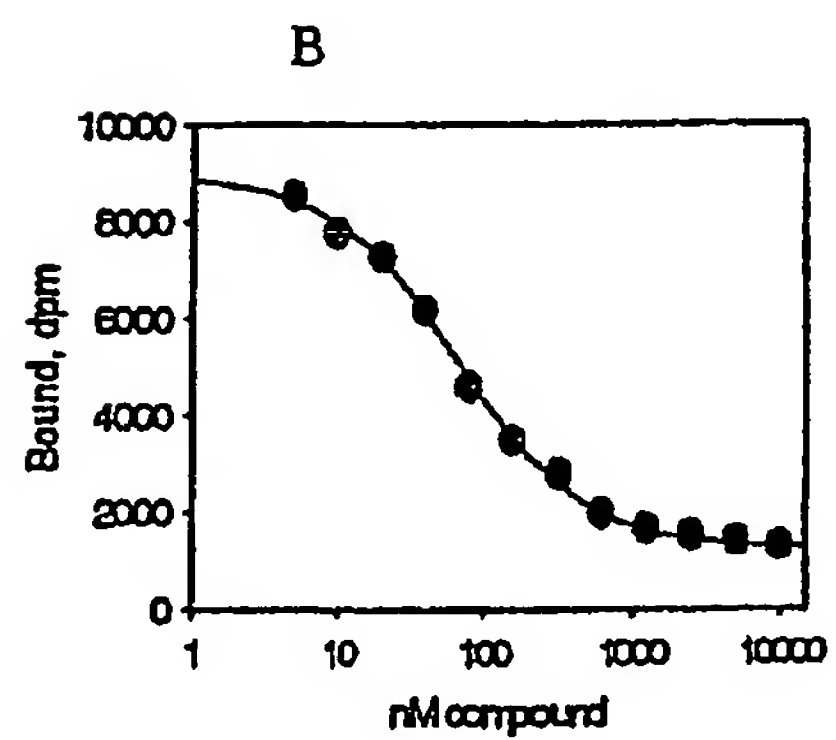
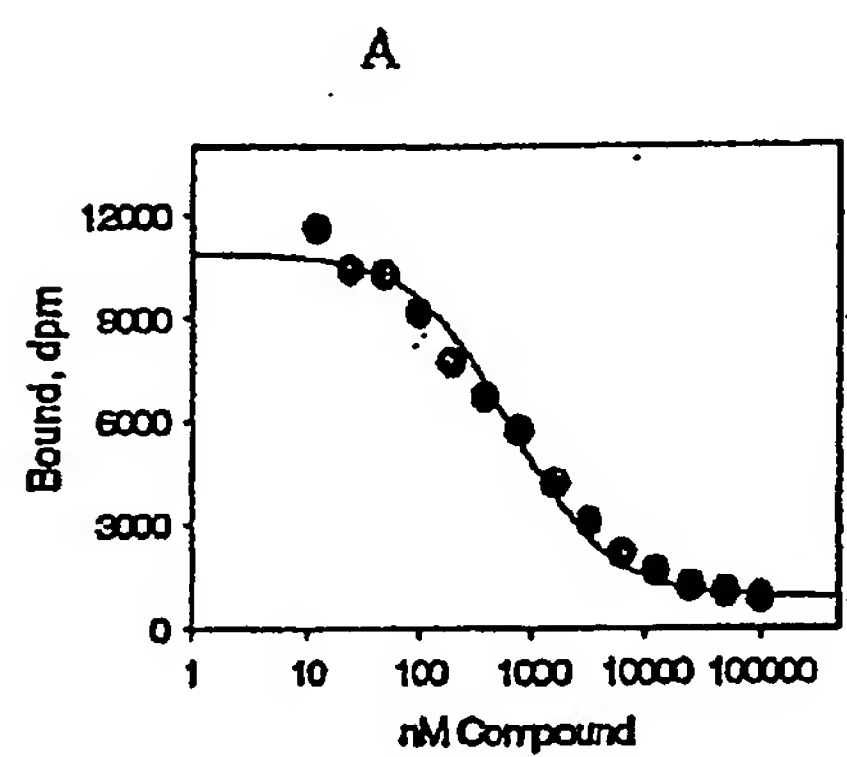
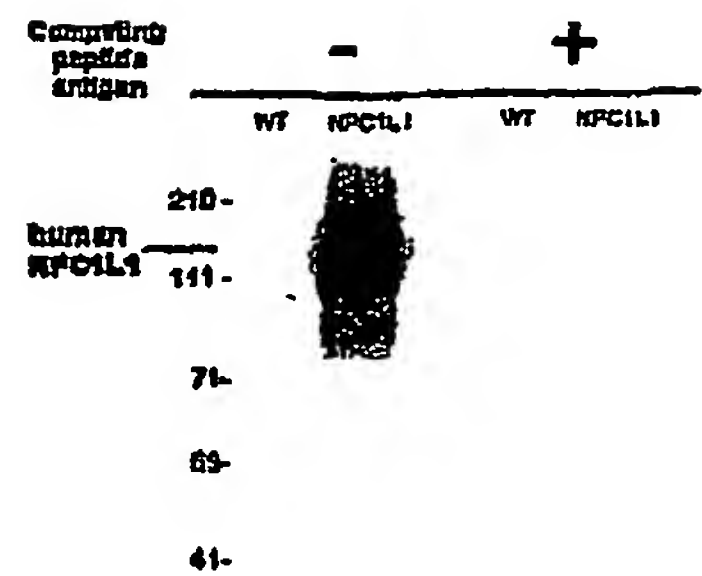


FIGURE 13

Panel 1



Panel 2

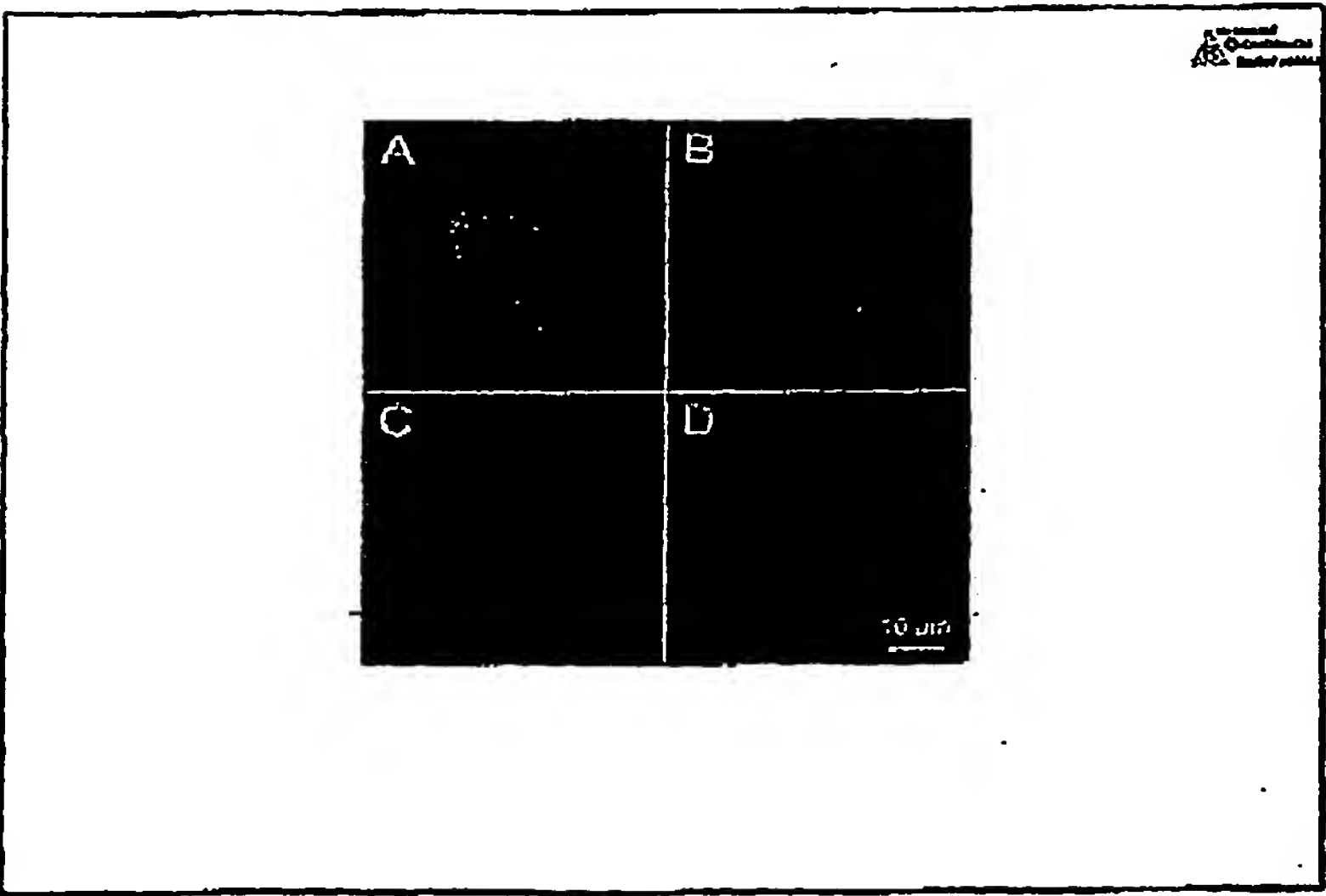


FIGURE 14

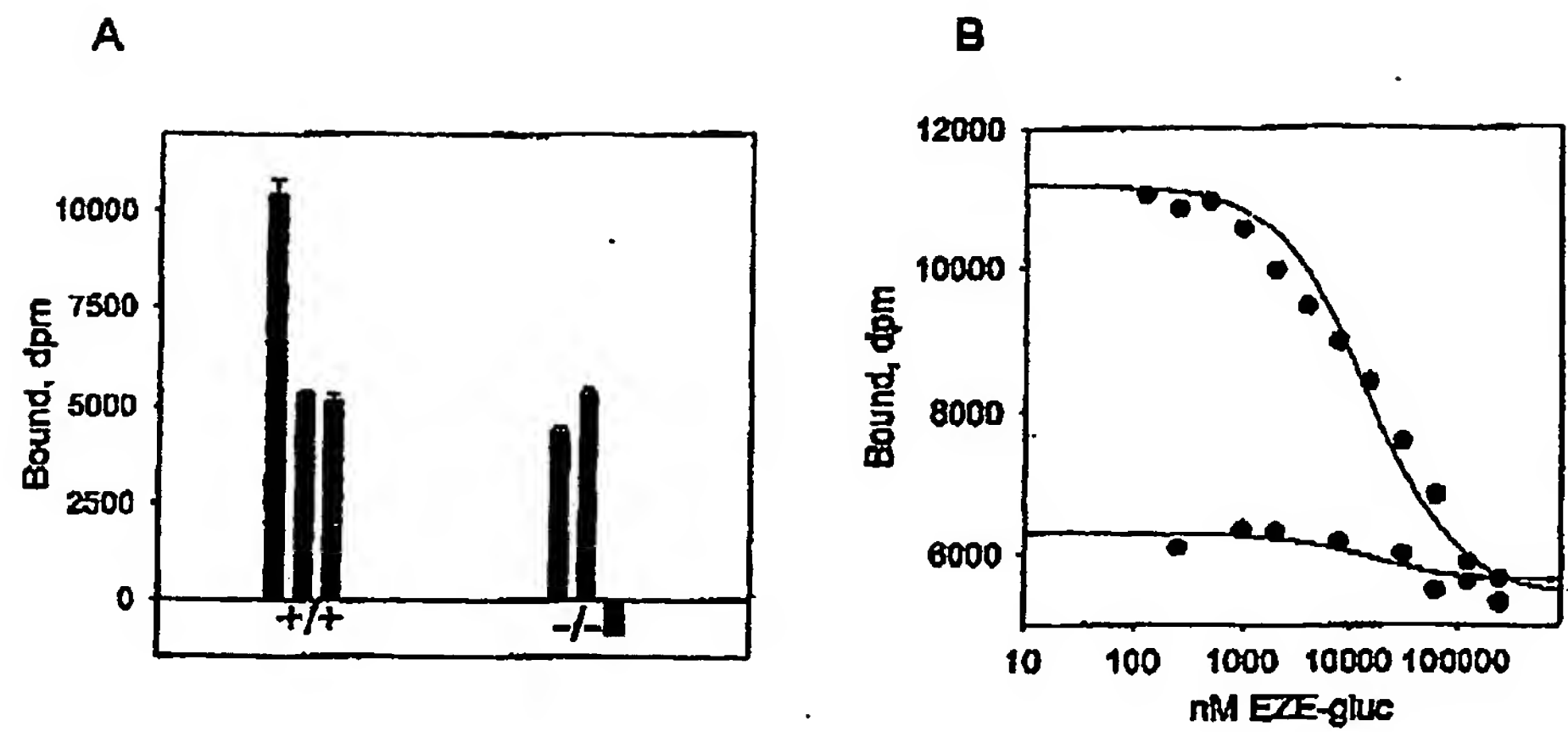
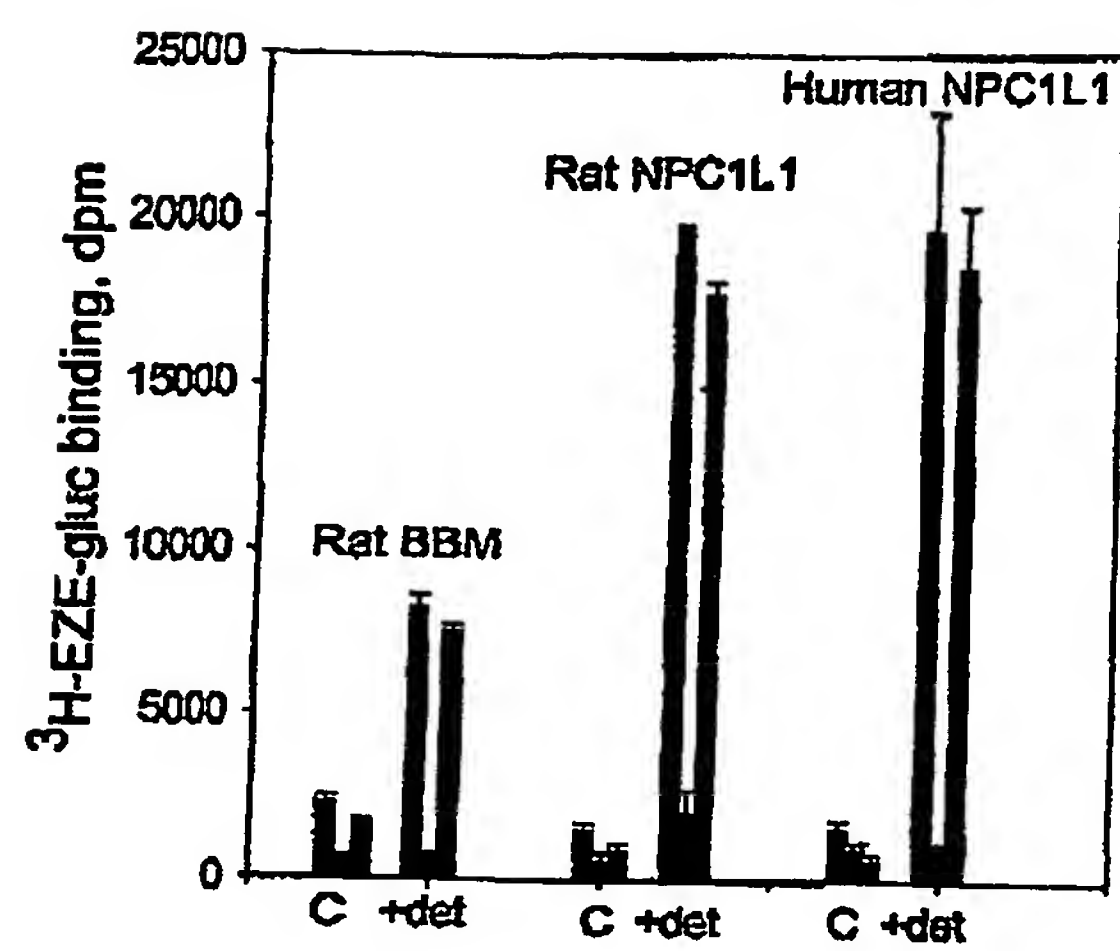


FIGURE 15



SEQUENCE LISTING

<110> Garcia-Calvo, Margarita

<120> NPC1L1 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF

<130> 36134-PCT 074669.0111

<140> To Be Assigned

<141> 2005-01-14

<150> US 60/537,341

<151> 2004-01-16

<160> 51

<170> PatentIn version 3.1

<210> 1

<211> 3996

<212> DNA

<213> Rattus sp.

<220>

<221> CDS

<222> (1) .. (3996)

<223>

<400> 1

atg gca gct gcc tgg ctg gga tgg ctg ctc tgg gcc ctg ctc ctg agc	48
Met Ala Ala Ala Trp Leu Gly Trp Leu Leu Trp Ala Leu Leu Leu Ser	
1 5 10 15	

gcg gcc cag ggt gag cta tac aca ccc aaa cac gaa gct ggg gtc tgc	96
Ala Ala Gln Gly Glu Leu Tyr Thr Pro Lys His Glu Ala Gly Val Cys	
20 25 30	

acc ttt tac gaa gag tgc ggg aaa aac cca gag ctc tct gga ggc ctc	144
Thr Phe Tyr Glu Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Gly Leu	

35	40	45	
acg tca cta tcc aat gta tcc tgc ctg tct aac acc ccg gcc cgc cac Thr Ser Leu Ser Asn Val Ser Cys Leu Ser Asn Thr Pro Ala Arg His 50 55 60			192
gtc acg ggt gaa cac ctg gct ctt ctc cag cgc atc tgt ccc cgc ctg Val Thr Gly Glu His Leu Ala Leu Leu Gln Arg Ile Cys Pro Arg Leu 65 70 75 80			240
tac aac ggc ccc aat acc act ttt gcc tgt tgc tct acc aag cag ctg Tyr Asn Gly Pro Asn Thr Thr Phe Ala Cys Cys Ser Thr Lys Gln Leu 85 90 95			288
ctg tcc tta gaa agc agc atg tcc atc acc aag gcc ctt ctc acg cgc Leu Ser Leu Glu Ser Ser Met Ser Ile Thr Lys Ala Leu Leu Thr Arg 100 105 110			336
tgc ccg gcc tgc tct gac aat ttt gtg agc tta cac tgc cac aac act Cys Pro Ala Cys Ser Asp Asn Phe Val Ser Leu His Cys His Asn Thr 115 120 125			384
tgc agc cct gac cag agc ctc ttc atc aac gtc acc ccg gtg gtt gag Cys Ser Pro Asp Gln Ser Leu Phe Ile Asn Val Thr Arg Val Val Glu 130 135 140			432
cgg ggc gct gga gag cct cct gcc gtg gtg gcc tat gag gcc ttt tat Arg Gly Ala Gly Glu Pro Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr 145 150 155 160			480
cag cgc agc ttt gct gag aag gcc tat gag tcc tgc agc cag gtg cgc Gln Arg Ser Phe Ala Glu Lys Ala Tyr Glu Ser Cys Ser Gln Val Arg 165 170 175			528
atc cct gcg gcc gct tcc ttg gcc gtg ggc agc atg tgt gga gtg tat Ile Pro Ala Ala Ala Ser Leu Ala Val Gly Ser Met Cys Gly Val Tyr 180 185 190			576
ggc tcc gcc ctc tgc aat gct cag cgc tgg ctc aac ttc caa gga gac Gly Ser Ala Leu Cys Asn Ala Gln Arg Trp Leu Asn Phe Gln Gly Asp 195 200 205			624
aca ggg aat ggc ctg gct ccg ctg gat atc acc ttc cac ctc ttg gag Thr Gly Asn Gly Leu Ala Pro Leu Asp Ile Thr Phe His Leu Leu Glu 210 215 220			672
cct ggc cag gcc cta ccg gat ggg atc cag cca ctg aat ggg aag atc Pro Gly Gln Ala Leu Pro Asp Gly Ile Gln Pro Leu Asn Gly Lys Ile 225 230 235 240			720
gca ccc tgc aac gag tct cag ggt gat gac tca gca gtc tgc tcc tgc Ala Pro Cys Asn Glu Ser Gln Gly Asp Asp Ser Ala Val Cys Ser Cys 245 250 255			768
cag gac tgt gcg gcg tcc tgc cct gtc atc cct ccg ccc gag gcc ttg Gln Asp Cys Ala Ala Ser Cys Pro Val Ile Pro Pro Pro Glu Ala Leu 260 265 270			816
cgc cct tcc ttc tac atg ggt cgc atg cca ggc tgg ctg gcc ctc atc			864

Arg	Pro	Ser	Phe	Tyr	Met	Gly	Arg	Met	Pro	Gly	Trp	Leu	Ala	Leu	Ile		
	275						280					285					
atc	atc	ttc	act	gct	gtc	ttt	gtg	ttg	ctc	tct	gca	gtc	ctt	gtg	cgt	912	
Ile	Ile	Phe	Thr	Ala	Val	Phe	Val	Leu	Leu	Ser	Ala	Val	Leu	Val	Arg		
	290					295					300						
ctc	cga	gtg	gtt	tcc	aac	agg	aac	aag	aac	aag	gca	gaa	ggc	ccc	cag	960	
Leu	Arg	Val	Val	Ser	Asn	Arg	Asn	Lys	Asn	Lys	Ala	Glu	Gly	Pro	Gln		
	305				310					315					320		
gaa	gcc	ccc	aaa	ctc	cct	cat	aag	cac	aaa	ctc	tca	ccc	cat	acc	atc	1008	
Glu	Ala	Pro	Lys	Leu	Pro	His	Lys	His	Lys	Leu	Ser	Pro	His	Thr	Ile		
				325					330					335			
ctg	ggc	cgg	ttc	ttc	cag	aac	tgg	ggc	aca	agg	gtg	gcc	tcg	tgg	cca	1056	
Leu	Gly	Arg	Phe	Phe	Gln	Asn	Trp	Gly	Thr	Arg	Val	Ala	Ser	Trp	Pro		
			340					345					350				
ctc	acc	gtc	tta	gca	ctg	tcc	ttc	atc	gtt	gtg	ata	gcc	tta	gca	gca	1104	
Leu	Thr	Val	Leu	Ala	Leu	Ser	Phe	Ile	Val	Val	Ile	Ala	Leu	Ala	Ala		
		355					360					365					
ggc	ctg	acc	ttt	att	gaa	ctc	acc	aca	gac	cct	gtg	gaa	ctg	tgg	tcg	1152	
Gly	Leu	Thr	Phe	Ile	Glu	Leu	Thr	Thr	Asp	Pro	Val	Glu	Leu	Trp	Ser		
	370					375					380						
gcc	ccc	aag	agc	cag	gcc	cgg	aaa	gag	aag	tct	ttc	cat	gat	gag	cat	1200	
Ala	Pro	Lys	Ser	Gln	Ala	Arg	Lys	Glu	Lys	Ser	Phe	His	Asp	Glu	His		
	385				390					395					400		
ttc	ggc	ccc	ttc	ttt	cga	acc	aac	cag	att	ttc	gtg	aca	gct	cgg	aac	1248	
Phe	Gly	Pro	Phe	Phe	Arg	Thr	Asn	Gln	Ile	Phe	Val	Thr	Ala	Arg	Asn		
				405				410						415			
agg	tcc	agc	tac	aag	tac	gac	tcc	cta	ctg	cta	ggg	tcc	aag	aac	ttc	1296	
Arg	Ser	Ser	Tyr	Lys	Tyr	Asp	Ser	Leu	Leu	Leu	Gly	Ser	Lys	Asn	Phe		
			420					425					430				
agt	ggg	atc	ctg	tcc	ctg	gac	ttc	ctg	ctg	gag	ctg	ctg	gag	ctt	cag	1344	
Ser	Gly	Ile	Leu	Ser	Leu	Asp	Phe	Leu	Leu	Glu	Leu	Leu	Glu	Leu	Gln		
		435				440						445					
gag	agg	ctt	cga	cac	ctg	caa	gtg	tgg	tcc	cct	gag	gca	gag	cgc	aac	1392	
Glu	Arg	Leu	Arg	His	Leu	Gln	Val	Trp	Ser	Pro	Glu	Ala	Glu	Arg	Asn		
	450					455					460						
atc	tcc	ctc	cag	gac	atc	tgc	tat	gcc	ccc	ctc	aac	cca	tat	aac	acc	1440	
Ile	Ser	Leu	Gln	Asp	Ile	Cys	Tyr	Ala	Pro	Leu	Asn	Pro	Tyr	Asn	Thr		
	465				470					475					480		
agc	ctc	tcc	gac	tgc	tgt	gtc	aac	agc	ctc	ctt	cag	tac	ttc	cag	aac	1488	
Ser	Leu	Ser	Asp	Cys	Cys	Val	Asn	Ser	Leu	Leu	Gln	Tyr	Phe	Gln	Asn		
			485					490						495			
aac	cgc	acc	ctc	ctg	atg	ctc	acg	gcc	aac	cag	act	ctg	aat	ggc	cag	1536	
Asn	Arg	Thr	Leu	Leu	Met	Leu	Thr	Ala	Asn	Gln	Thr	Leu	Asn	Gly	Gln		
		500						505					510				

acc tcc ctg gtg gac tgg aag gac cat ttc ctc tac tgt gca aat gcc Thr Ser Leu Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala 515 520 525	1584
cct ctc acg ttc aaa gat ggc acg tct ctg gcc ctg agc tgc atg gct Pro Leu Thr Phe Lys Asp Gly Thr Ser Leu Ala Leu Ser Cys Met Ala 530 535 540	1632
gac tac ggg gct cct gtc ttc ccc ttc ctt gct gtt ggg gga tac caa Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Val Gly Gly Tyr Gln 545 550 555 560	1680
ggc acg gac tat tcc gag gca gaa gcg ctg atc ata acc ttc tct ctc Gly Thr Asp Tyr Ser Glu Ala Glu Ala Leu Ile Ile Thr Phe Ser Leu 565 570 575	1728
aat aac tac ccc gct gat gat ccc cgc atg gcc cag gcc aag ctc tgg Asn Asn Tyr Pro Ala Asp Asp Pro Arg Met Ala Gln Ala Lys Leu Trp 580 585 590	1776
gag gag gct ttc ttg aag gaa atg gaa tcc ttc cag agg aac aca agt Glu Glu Ala Phe Leu Lys Glu Met Glu Ser Phe Gln Arg Asn Thr Ser 595 600 605	1824
gac aag ttc cag gtt gcg ttc tca gct gag cgc tct ctg gag gat gag Asp Lys Phe Gln Val Ala Phe Ser Ala Glu Arg Ser Leu Glu Asp Glu 610 615 620	1872
atc aac cgc acc acc atc cag gac ctg cct gtc ttt gcc gtc agc tac Ile Asn Arg Thr Thr Ile Gln Asp Leu Pro Val Phe Ala Val Ser Tyr 625 630 635 640	1920
att atc gtc ttc ctg tac atc tcc ctg gcc ctg ggc agc tac tcc aga Ile Ile Val Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Arg 645 650 655	1968
tgc agc cga gta gcg gtg gag tcc aag gct act ctg ggc cta ggt ggg Cys Ser Arg Val Ala Val Glu Ser Lys Ala Thr Leu Gly Leu Gly Gly 660 665 670	2016
gtg att gtt gtg ctg gga gca gtt ctg gct gcc atg ggc ttc tac tcc Val Ile Val Val Leu Gly Ala Val Leu Ala Ala Met Gly Phe Tyr Ser 675 680 685	2064
tac ctg ggt gtc ccc tct tct ctg gtt atc atc caa gtg gta cct ttc Tyr Leu Gly Val Pro Ser Ser Leu Val Ile Ile Gln Val Val Pro Phe 690 695 700	2112
ctg gtg cta gct gtg gga gct gac aac atc ttc atc ttt gtt ctt gag Leu Val Leu Ala Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu 705 710 715 720	2160
tac cag agg cta cct agg atg cct ggg gaa cag cga gag gct cac att Tyr Gln Arg Leu Pro Arg Met Pro Gly Glu Gln Arg Glu Ala His Ile 725 730 735	2208
ggc cgc acc ctg ggc agt gtg gcc ccc agc atg ctg ctg tgc agc ctc Gly Arg Thr Leu Gly Ser Val Ala Pro Ser Met Leu Leu Cys Ser Leu 740 745 750	2256

tct gag gcc atc tgc ttc ttt cta ggg gcc ctg acc ccc atg cca gct Ser Glu Ala Ile Cys Phe Phe Leu Gly Ala Leu Thr Pro Met Pro Ala 755 760 765	2304
gtg agg acc ttc gcc ttg acc tct ggc tta gca att atc ctc gac ttc Val Arg Thr Phe Ala Leu Thr Ser Gly Leu Ala Ile Ile Leu Asp Phe 770 775 780	2352
ctg ctc cag atg act gcc ttt gtg gcc ctg ctc tcc ctg gat agc aag Leu Leu Gln Met Thr Ala Phe Val Ala Leu Leu Ser Leu Asp Ser Lys 785 790 795 800	2400
agg cag gag gcc tct cgc ccg gat gtc tta tgc tgc ttt tca acc cgg Arg Gln Glu Ala Ser Arg Pro Asp Val Leu Cys Cys Phe Ser Thr Arg 805 810 815	2448
aag ctg ccc cca cct aaa gaa aaa gaa ggc ctc tta ctc cgc ttc ttc Lys Leu Pro Pro Pro Lys Glu Lys Glu Gly Leu Leu Leu Arg Phe Phe 820 825 830	2496
cgc aag ata tac gct cct ttc ctg ctg cac aga ttc atc cgc cct gtt Arg Lys Ile Tyr Ala Pro Phe Leu Leu His Arg Phe Ile Arg Pro Val 835 840 845	2544
gtg atg ctg ctg ttt ctg acc ctg ttt gga gca aat ctc tac tta atg Val Met Leu Leu Phe Leu Thr Leu Phe Gly Ala Asn Leu Tyr Leu Met 850 855 860	2592
tgc aac atc aac gtg ggg cta gac cag gag ctg gct ctg ccc aag gac Cys Asn Ile Asn Val Gly Leu Asp Gln Glu Leu Ala Leu Pro Lys Asp 865 870 875 880	2640
tcg tac ttg ata gac tac ttc ctc ttt ctg aac cga tac ctt gaa gtg Ser Tyr Leu Ile Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Leu Glu Val 885 890 895	2688
ggg cct cca gtg tac ttt gtc acc acc tcg ggc ttc aac ttc tcc agc Gly Pro Pro Val Tyr Phe Val Thr Thr Ser Gly Phe Asn Phe Ser Ser 900 905 910	2736
gag gca ggc atg aac gcc act tgc tct agc gca ggc tgt aag agc ttc Glu Ala Gly Met Asn Ala Thr Cys Ser Ser Ala Gly Cys Lys Ser Phe 915 920 925	2784
tcc cta acc cag aaa atc cag tat gcc agt gaa ttc cct gac cag tct Ser Leu Thr Gln Lys Ile Gln Tyr Ala Ser Glu Phe Pro Asp Gln Ser 930 935 940	2832
tac gtg gct att gct gca tcc tcc tgg gta gat gac ttc atc gac tgg Tyr Val Ala Ile Ala Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp 945 950 955 960	2880
ctg acc ccg tcc tcc tcc tgc tgt cgc ctt tat ata cgt ggc ccc cat Leu Thr Pro Ser Ser Ser Cys Cys Arg Leu Tyr Ile Arg Gly Pro His 965 970 975	2928
aag gat gag ttc tgt ccc tca acg gat act tcc ttc aac tgc tta aaa Lys Asp Glu Phe Cys Pro Ser Thr Asp Thr Ser Phe Asn Cys Leu Lys	2976

980					985					990					
aac tgc	atg aac	cgc act	ctg ggt	cct gtg	agg ccc	aca gcg	gaa cag								3024
Asn Cys	Met Asn	Arg Thr	Leu Gly	Pro Val	Arg Pro	Thr Ala	Glu Gln								
	995		1000			1005									
ttt cat	aag tac	ctg ccc	tgg ttc	ctg aat	gat ccg	ccc aat	atc								3069
Phe His	Lys Tyr	Leu Pro	Trp Phe	Leu Asn	Asp Pro	Pro Asn	Ile								
	1010		1015		1020										
aga tgt	ccc aaa	ggg ggt	cta gca	gcg tat	aga acg	tct gtg	aat								3114
Arg Cys	Pro Lys	Gly Gly	Leu Ala	Ala Tyr	Arg Thr	Ser Val	Asn								
	1025		1030		1035										
ttg agc	tca gat	ggc cag	gtt ata	gcc tcc	cag ttc	atg gcc	tac								3159
Leu Ser	Ser Asp	Gly Gln	Val Ile	Ala Ser	Gln Phe	Met Ala	Tyr								
	1040		1045		1050										
cac aag	ccc tta	agg aac	tca cag	gac ttc	aca gaa	gct ctc	cgg								3204
His Lys	Pro Leu	Arg Asn	Ser Gln	Asp Phe	Thr Glu	Ala Leu	Arg								
	1055		1060		1065										
gcg tcc	cgg ttg	cta gca	gcc aac	atc aca	gct gac	cta cgg	aag								3249
Ala Ser	Arg Leu	Leu Ala	Ala Asn	Ile Thr	Ala Asp	Leu Arg	Lys								
	1070		1075		1080										
gtg cct	ggg aca	gat cca	aac ttt	gag gtc	ttc cct	tac acg	atc								3294
Val Pro	Gly Thr	Asp Pro	Asn Phe	Glu Val	Phe Pro	Tyr Thr	Ile								
	1085		1090		1095										
tcc aac	gtg ttc	tac cag	caa tac	ctg acg	gtc ctt	cct gag	gga								3339
Ser Asn	Val Phe	Tyr Gln	Gln Tyr	Leu Thr	Val Leu	Pro Glu	Gly								
	1100		1105		1110										
atc ttc	acc ctt	gct ctt	tgc ttt	gtg ccc	acc ttt	gtt gtc	tgc								3384
Ile Phe	Thr Leu	Ala Leu	Cys Phe	Val Pro	Thr Phe	Val Val	Cys								
	1115		1120		1125										
tac ctc	cta ctg	ggc ctg	gac atg	tgc tca	ggg atc	ctc aac	cta								3429
Tyr Leu	Leu Leu	Gly Leu	Asp Met	Cys Ser	Gly Ile	Leu Asn	Leu								
	1130		1135		1140										
ctc tcc	atc att	atg att	ctc gtg	gac acc	att ggc	ctc atg	gct								3474
Leu Ser	Ile Ile	Met Ile	Leu Val	Asp Thr	Ile Gly	Leu Met	Ala								
	1145		1150		1155										
gtg tgg	ggg atc	agc tat	aat gcg	gta tcc	ctc atc	aac ctt	gtc								3519
Val Trp	Gly Ile	Ser Tyr	Asn Ala	Val Ser	Leu Ile	Asn Leu	Val								
	1160		1165		1170										
acg gca	gtg ggc	atg tct	gtg gag	ttt gtg	tcc cac	atc act	cgg								3564
Thr Ala	Val Gly	Met Ser	Val Glu	Phe Val	Ser His	Ile Thr	Arg								
	1175		1180		1185										
tcc ttt	gct gta	agc acc	aag cct	acc cgg	ctg gag	agg gct	aaa								3609
Ser Phe	Ala Val	Ser Thr	Lys Pro	Thr Arg	Leu Glu	Arg Ala	Lys								
	1190		1195		1200										
gat gct	act gtc	ttc atg	ggc agt	gcg gtg	ttt gct	gga gtg	gcc								3654

Asp	Ala	Thr	Val	Phe	Met	Gly	Ser	Ala	Val	Phe	Ala	Gly	Val	Ala		
1205						1210					1215					
atg	acc	aac	ttc	cca	ggc	atc	ctc	atc	ttg	ggc	ttt	gcc	caa	gcc	3699	
Met	Thr	Asn	Phe	Pro	Gly	Ile	Leu	Ile	Leu	Gly	Phe	Ala	Gln	Ala		
1220						1225					1230					
cag	ctt	att	cag	atc	ttc	ttc	ttc	cgc	ctc	aac	ctt	ctg	atc	acc	3744	
Gln	Leu	Ile	Gln	Ile	Phe	Phe	Phe	Arg	Leu	Asn	Leu	Leu	Ile	Thr		
1235						1240					1245					
ttg	ctg	ggt	ctg	ctg	cat	ggc	ctg	gtc	ttc	ctg	ccg	gtt	gtc	ctc	3789	
Leu	Leu	Gly	Leu	Leu	His	Gly	Leu	Val	Phe	Leu	Pro	Val	Val	Leu		
1250						1255					1260					
agc	tat	ctg	gga	cca	gat	gtt	aac	caa	gct	ctg	gta	cag	gag	gag	3834	
Ser	Tyr	Leu	Gly	Pro	Asp	Val	Asn	Gln	Ala	Leu	Val	Gln	Glu	Glu		
1265						1270					1275					
aaa	cta	gcc	agc	gag	gca	gca	gtg	gcc	cca	gag	cct	tct	tgc	cca	3879	
Lys	Leu	Ala	Ser	Glu	Ala	Ala	Val	Ala	Pro	Glu	Pro	Ser	Cys	Pro		
1280						1285					1290					
cag	tac	ccc	tcc	cct	gct	gat	gcg	gat	gcc	aat	gtt	aac	tac	ggc	3924	
Gln	Tyr	Pro	Ser	Pro	Ala	Asp	Ala	Asp	Ala	Asn	Val	Asn	Tyr	Gly		
1295						1300					1305					
ttt	gcc	cca	gaa	ctt	gcc	cac	gga	gct	aat	gct	gct	aga	agc	tct	3969	
Phe	Ala	Pro	Glu	Leu	Ala	His	Gly	Ala	Asn	Ala	Ala	Arg	Ser	Ser		
1310						1315					1320					
ttg	ccc	aaa	agt	gac	caa	aag	ttc	taa							3996	
Leu	Pro	Lys	Ser	Asp	Gln	Lys	Phe									
1325						1330										

<210> 2
 <211> 1331
 <212> PRT
 <213> Rattus sp.

<400> 2

Met	Ala	Ala	Ala	Trp	Leu	Gly	Trp	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Ser
1				5				10						15	
Ala	Ala	Gln	Gly	Glu	Leu	Tyr	Thr	Pro	Lys	His	Glu	Ala	Gly	Val	Cys
		20					25						30		
Thr	Phe	Tyr	Glu	Glu	Cys	Gly	Lys	Asn	Pro	Glu	Leu	Ser	Gly	Gly	Leu
	35					40						45			

Thr Ser Leu Ser Asn Val Ser Cys Leu Ser Asn Thr Pro Ala Arg His
50 55 60

Val Thr Gly Glu His Leu Ala Leu Leu Gln Arg Ile Cys Pro Arg Leu
65 70 75 80

Tyr Asn Gly Pro Asn Thr Thr Phe Ala Cys Cys Ser Thr Lys Gln Leu
85 90 95

Leu Ser Leu Glu Ser Ser Met Ser Ile Thr Lys Ala Leu Leu Thr Arg
100 105 110

Cys Pro Ala Cys Ser Asp Asn Phe Val Ser Leu His Cys His Asn Thr
115 120 125

Cys Ser Pro Asp Gln Ser Leu Phe Ile Asn Val Thr Arg Val Val Glu
130 135 140

Arg Gly Ala Gly Glu Pro Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr
145 150 155 160

Gln Arg Ser Phe Ala Glu Lys Ala Tyr Glu Ser Cys Ser Gln Val Arg
165 170 175

Ile Pro Ala Ala Ala Ser Leu Ala Val Gly Ser Met Cys Gly Val Tyr
180 185 190

Gly Ser Ala Leu Cys Asn Ala Gln Arg Trp Leu Asn Phe Gln Gly Asp
195 200 205

Thr Gly Asn Gly Leu Ala Pro Leu Asp Ile Thr Phe His Leu Leu Glu
210 215 220

Pro Gly Gln Ala Leu Pro Asp Gly Ile Gln Pro Leu Asn Gly Lys Ile
225 230 235 240

Ala Pro Cys Asn Glu Ser Gln Gly Asp Asp Ser Ala Val Cys Ser Cys
245 250 255

Gln Asp Cys Ala Ala Ser Cys Pro Val Ile Pro Pro Pro Glu Ala Leu
260 265 270

Arg Pro Ser Phe Tyr Met Gly Arg Met Pro Gly Trp Leu Ala Leu Ile
275 280 285

Ile Ile Phe Thr Ala Val Phe Val Leu Leu Ser Ala Val Leu Val Arg
290 295 300

Leu Arg Val Val Ser Asn Arg Asn Lys Asn Lys Ala Glu Gly Pro Gln
305 310 315 320

Glu Ala Pro Lys Leu Pro His Lys His Lys Leu Ser Pro His Thr Ile
325 330 335

Leu Gly Arg Phe Phe Gln Asn Trp Gly Thr Arg Val Ala Ser Trp Pro
340 345 350

Leu Thr Val Leu Ala Leu Ser Phe Ile Val Val Ile Ala Leu Ala Ala
355 360 365

Gly Leu Thr Phe Ile Glu Leu Thr Thr Asp Pro Val Glu Leu Trp Ser
370 375 380

Ala Pro Lys Ser Gln Ala Arg Lys Glu Lys Ser Phe His Asp Glu His
385 390 395 400

Phe Gly Pro Phe Phe Arg Thr Asn Gln Ile Phe Val Thr Ala Arg Asn
405 410 415

Arg Ser Ser Tyr Lys Tyr Asp Ser Leu Leu Leu Gly Ser Lys Asn Phe
420 425 430

Ser Gly Ile Leu Ser Leu Asp Phe Leu Leu Glu Leu Leu Glu Leu Gln
435 440 445

Glu Arg Leu Arg His Leu Gln Val Trp Ser Pro Glu Ala Glu Arg Asn
450 455 460

Ile Ser Leu Gln Asp Ile Cys Tyr Ala Pro Leu Asn Pro Tyr Asn Thr
465 470 475 480

Ser Leu Ser Asp Cys Cys Val Asn Ser Leu Leu Gln Tyr Phe Gln Asn
485 490 495

Asn Arg Thr Leu Leu Met Leu Thr Ala Asn Gln Thr Leu Asn Gly Gln
500 505 510

Thr Ser Leu Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala

515	520	525
Pro Leu Thr Phe Lys Asp Gly Thr Ser Leu Ala Leu Ser Cys Met Ala 530 535 540		
Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Val Gly Gly Tyr Gln 545 550 555 560		
Gly Thr Asp Tyr Ser Glu Ala Glu Ala Leu Ile Ile Thr Phe Ser Leu 565 570 575		
Asn Asn Tyr Pro Ala Asp Asp Pro Arg Met Ala Gln Ala Lys Leu Trp 580 585 590		
Glu Glu Ala Phe Leu Lys Glu Met Glu Ser Phe Gln Arg Asn Thr Ser 595 600 605		
Asp Lys Phe Gln Val Ala Phe Ser Ala Glu Arg Ser Leu Glu Asp Glu 610 615 620		
Ile Asn Arg Thr Thr Ile Gln Asp Leu Pro Val Phe Ala Val Ser Tyr 625 630 635 640		
Ile Ile Val Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Arg 645 650 655		
Cys Ser Arg Val Ala Val Glu Ser Lys Ala Thr Leu Gly Leu Gly Gly 660 665 670		
Val Ile Val Val Leu Gly Ala Val Leu Ala Ala Met Gly Phe Tyr Ser 675 680 685		
Tyr Leu Gly Val Pro Ser Ser Leu Val Ile Ile Gln Val Val Pro Phe 690 695 700		
Leu Val Leu Ala Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu 705 710 715 720		
Tyr Gln Arg Leu Pro Arg Met Pro Gly Glu Gln Arg Glu Ala His Ile 725 730 735		
Gly Arg Thr Leu Gly Ser Val Ala Pro Ser Met Leu Leu Cys Ser Leu 740 745 750		

Ser Glu Ala Ile Cys Phe Phe Leu Gly Ala Leu Thr Pro Met Pro Ala
755 760 765

Val Arg Thr Phe Ala Leu Thr Ser Gly Leu Ala Ile Ile Leu Asp Phe
770 775 780

Leu Leu Gln Met Thr Ala Phe Val Ala Leu Leu Ser Leu Asp Ser Lys
785 790 795 800

Arg Gln Glu Ala Ser Arg Pro Asp Val Leu Cys Cys Phe Ser Thr Arg
805 810 815

Lys Leu Pro Pro Pro Lys Glu Lys Glu Gly Leu Leu Leu Arg Phe Phe
820 825 830

Arg Lys Ile Tyr Ala Pro Phe Leu Leu His Arg Phe Ile Arg Pro Val
835 840 845

Val Met Leu Leu Phe Leu Thr Leu Phe Gly Ala Asn Leu Tyr Leu Met
850 855 860

Cys Asn Ile Asn Val Gly Leu Asp Gln Glu Leu Ala Leu Pro Lys Asp
865 870 875 880

Ser Tyr Leu Ile Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Leu Glu Val
885 890 895

Gly Pro Pro Val Tyr Phe Val Thr Thr Ser Gly Phe Asn Phe Ser Ser
900 905 910

Glu Ala Gly Met Asn Ala Thr Cys Ser Ser Ala Gly Cys Lys Ser Phe
915 920 925

Ser Leu Thr Gln Lys Ile Gln Tyr Ala Ser Glu Phe Pro Asp Gln Ser
930 935 940

Tyr Val Ala Ile Ala Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp
945 950 955 960

Leu Thr Pro Ser Ser Ser Cys Cys Arg Leu Tyr Ile Arg Gly Pro His
965 970 975

Lys Asp Glu Phe Cys Pro Ser Thr Asp Thr Ser Phe Asn Cys Leu Lys
980 985 990

Asn Cys Met Asn Arg Thr Leu Gly Pro Val Arg Pro Thr Ala Glu Gln
995 1000 1005

Phe His Lys Tyr Leu Pro Trp Phe Leu Asn Asp Pro Pro Asn Ile
1010 1015 1020

Arg Cys Pro Lys Gly Gly Leu Ala Ala Tyr Arg Thr Ser Val Asn
1025 1030 1035

Leu Ser Ser Asp Gly Gln Val Ile Ala Ser Gln Phe Met Ala Tyr
1040 1045 1050

His Lys Pro Leu Arg Asn Ser Gln Asp Phe Thr Glu Ala Leu Arg
1055 1060 1065

Ala Ser Arg Leu Leu Ala Ala Asn Ile Thr Ala Asp Leu Arg Lys
1070 1075 1080

Val Pro Gly Thr Asp Pro Asn Phe Glu Val Phe Pro Tyr Thr Ile
1085 1090 1095

Ser Asn Val Phe Tyr Gln Gln Tyr Leu Thr Val Leu Pro Glu Gly
1100 1105 1110

Ile Phe Thr Leu Ala Leu Cys Phe Val Pro Thr Phe Val Val Cys
1115 1120 1125

Tyr Leu Leu Leu Gly Leu Asp Met Cys Ser Gly Ile Leu Asn Leu
1130 1135 1140

Leu Ser Ile Ile Met Ile Leu Val Asp Thr Ile Gly Leu Met Ala
1145 1150 1155

Val Trp Gly Ile Ser Tyr Asn Ala Val Ser Leu Ile Asn Leu Val
1160 1165 1170

Thr Ala Val Gly Met Ser Val Glu Phe Val Ser His Ile Thr Arg
1175 1180 1185

Ser Phe Ala Val Ser Thr Lys Pro Thr Arg Leu Glu Arg Ala Lys
1190 1195 1200

Asp Ala Thr Val Phe Met Gly Ser Ala Val Phe Ala Gly Val Ala
1205 1210 1215

Met Thr Asn Phe Pro Gly Ile Leu Ile Leu Gly Phe Ala Gln Ala
1220 1225 1230

Gln Leu Ile Gln Ile Phe Phe Phe Arg Leu Asn Leu Leu Ile Thr
1235 1240 1245

Leu Leu Gly Leu Leu His Gly Leu Val Phe Leu Pro Val Val Leu
1250 1255 1260

Ser Tyr Leu Gly Pro Asp Val Asn Gln Ala Leu Val Gln Glu Glu
1265 1270 1275

Lys Leu Ala Ser Glu Ala Ala Val Ala Pro Glu Pro Ser Cys Pro
1280 1285 1290

Gln Tyr Pro Ser Pro Ala Asp Ala Asp Ala Asn Val Asn Tyr Gly
1295 1300 1305

Phe Ala Pro Glu Leu Ala His Gly Ala Asn Ala Ala Arg Ser Ser
1310 1315 1320

Leu Pro Lys Ser Asp Gln Lys Phe
1325 1330

<210> 3

<211> 3999

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1) .. (3999)

<223>

<400> 3

atg gcg gag gcc ggc ctg agg ggc tgg ctg ctg tgg gcc ctg ctc ctg
Met Ala Glu Ala Gly Leu Arg Gly Trp Leu Leu Trp Ala Leu Leu Leu
1 5 10 15

48

cgc ttg gcc cag agt gag cct tac aca acc atc cac cag cct ggc tac Arg Leu Ala Gln Ser Glu Pro Tyr Thr Thr Ile His Gln Pro Gly Tyr 20 25 30	96
tgc gcc ttc tat gac gaa tgt ggg aag aac cca gag ctg tct gga agc Cys Ala Phe Tyr Asp Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Ser 35 40 45	144
ctc atg aca ctc tcc aac gtg tcc tgc ctg tcc aac acg ccg gcc cgc Leu Met Thr Leu Ser Asn Val Ser Cys Leu Ser Asn Thr Pro Ala Arg 50 55 60	192
aag atc aca ggt gat cac ctg atc cta tta cag aag atc tgc ccc cgc Lys Ile Thr Gly Asp His Leu Ile Leu Leu Gln Lys Ile Cys Pro Arg 65 70 75 80	240
ctc tac acc ggc ccc aac acc caa gcc tgc tgc tcc gcc aag cag ctg Leu Tyr Thr Gly Pro Asn Thr Gln Ala Cys Cys Ser Ala Lys Gln Leu 85 90 95	288
gta tca ctg gaa gcg agt ctg tgc atc acc aag gcc ctc ctc acc cgc Val Ser Leu Glu Ala Ser Leu Ser Ile Thr Lys Ala Leu Leu Thr Arg 100 105 110	336
tgc cca gcc tgc tct gac aat ttt gtg aac ctg cac tgc cac aac acg Cys Pro Ala Cys Ser Asp Asn Phe Val Asn Leu His Cys His Asn Thr 115 120 125	384
tgc agc ccc aat cag agc ctc ttc atc aat gtg acc cgc gtg gcc cag Cys Ser Pro Asn Gln Ser Leu Phe Ile Asn Val Thr Arg Val Ala Gln 130 135 140	432
cta ggg gct gga caa ctc cca gct gtg gtg gcc tat gag gcc ttc tac Leu Gly Ala Gly Gln Leu Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr 145 150 155 160	480
cag cat agc ttt gcc gag cag agc tat gac tcc tgc agc cgt gtg cgc Gln His Ser Phe Ala Glu Gln Ser Tyr Asp Ser Cys Ser Arg Val Arg 165 170 175	528
gtc cct gca gct gcc acg ctg gct gtg ggc acc atg tgt ggc gtg tat Val Pro Ala Ala Ala Thr Leu Ala Val Gly Thr Met Cys Gly Val Tyr 180 185 190	576
ggc tct gcc ctt tgc aat gcc cag cgc tgg ctc aac ttc cag gga gac Gly Ser Ala Leu Cys Asn Ala Gln Arg Trp Leu Asn Phe Gln Gly Asp 195 200 205	624
aca ggc aat ggt ctg gcc cca ctg gac atc acc ttc cac ctc ttg gag Thr Gly Asn Gly Leu Ala Pro Leu Asp Ile Thr Phe His Leu Leu Glu 210 215 220	672
cct ggc cag gcc gtg ggg agt ggg att cag cct ctg aat gag ggg gtt Pro Gly Gln Ala Val Gly Ser Gly Ile Gln Pro Leu Asn Glu Gly Val 225 230 235 240	720
gca cgt tgc aat gag tcc caa ggt gac gac gtg gcg acc tgc tcc tgc Ala Arg Cys Asn Glu Ser Gln Gly Asp Asp Val Ala Thr Cys Ser Cys 245 250 255	768

caa gac tgt gct gca tcc tgt cct gcc ata gcc cgc ccc cag gcc ctc Gln Asp Cys Ala Ala Ser Cys Pro Ala Ile Ala Arg Pro Gln Ala Leu 260 265 270	816
gac tcc acc ttc tac ctg ggc cag atg ccg ggc agt ctg gtc ctc atc Asp Ser Thr Phe Tyr Leu Gly Gln Met Pro Gly Ser Leu Val Leu Ile 275 280 285	864
atc atc ctc tgc tct gtc ttc gct gtg gtc acc atc ctg ctt gtg gga Ile Ile Leu Cys Ser Val Phe Ala Val Val Thr Ile Leu Leu Val Gly 290 295 300	912
ttc cgt gtg gcc ccc gcc agg gac aaa agc aag atg gtg gac ccc aag Phe Arg Val Ala Pro Ala Arg Asp Lys Ser Lys Met Val Asp Pro Lys 305 310 315 320	960
aag ggc acc agc ctc tct gac aag ctc agc ttc tcc acc cac acc ctc Lys Gly Thr Ser Leu Ser Asp Lys Leu Ser Phe Ser Thr His Thr Leu 325 330 335	1008
ctt ggc cag ttc ttc cag ggc tgg ggc acg tgg gtg gct tcg tgg cct Leu Gly Gln Phe Phe Gln Gly Trp Gly Thr Trp Val Ala Ser Trp Pro 340 345 350	1056
ctg acc atc ttg gtg cta tct gtc atc ccg gtg gtg gcc ttg gca gcg Leu Thr Ile Leu Val Leu Ser Val Ile Pro Val Val Ala Leu Ala Ala 355 360 365	1104
ggc ctg gtc ttt aca gaa ctc act acg gac ccc gtg gag ctg tgg tcg Gly Leu Val Phe Thr Glu Leu Thr Thr Asp Pro Val Glu Leu Trp Ser 370 375 380	1152
gcc ccc aac agc caa gcc cgg agt gag aaa gct ttc cat gac cag cat Ala Pro Asn Ser Gln Ala Arg Ser Glu Lys Ala Phe His Asp Gln His 385 390 395 400	1200
ttc ggc ccc ttc ttc cga acc aac cag gtg atc ctg acg gct cct aac Phe Gly Pro Phe Phe Arg Thr Asn Gln Val Ile Leu Thr Ala Pro Asn 405 410 415	1248
cgg tcc agc tac agg tat gac tct ctg ctg ctg ggg ccc aag aac ttc Arg Ser Ser Tyr Arg Tyr Asp Ser Leu Leu Leu Gly Pro Lys Asn Phe 420 425 430	1296
agc gga atc ctg gac ctg gac ttg ctg ctg gag ctg cta gag ctg cag Ser Gly Ile Leu Asp Leu Asp Leu Leu Leu Glu Leu Leu Glu Leu Gln 435 440 445	1344
gag agg ctg cgg cac ctc cag gta tgg tcg ccc gaa gca cag cgc aac Glu Arg Leu Arg His Leu Gln Val Trp Ser Pro Glu Ala Gln Arg Asn 450 455 460	1392
atc tcc ctg cag gac atc tgc tac gcc ccc ctc aat ccg gac aat acc Ile Ser Leu Gln Asp Ile Cys Tyr Ala Pro Leu Asn Pro Asp Asn Thr 465 470 475 480	1440
agt ctc tac gac tgc tgc atc aac agc ctc ctg cag tat ttc cag aac Ser Leu Tyr Asp Cys Cys Ile Asn Ser Leu Leu Gln Tyr Phe Gln Asn	1488

485	490	495	
aac cgc acg ctc ctg ctg ctc aca gcc aac cag aca ctg atg ggg cag Asn Arg Thr Leu Leu Leu Leu Thr Ala Asn Gln Thr Leu Met Gly Gln 500 505 510			1536
acc tcc caa gtc gac tgg aag gac cat ttt ctg tac tgt gcc aat gcc Thr Ser Gln Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala 515 520 525			1584
ccg ctc acc ttc aag gat ggc aca gcc ctg gcc ctg agc tgc atg gct Pro Leu Thr Phe Lys Asp Gly Thr Ala Leu Ala Leu Ser Cys Met Ala 530 535 540			1632
gac tac ggg gcc cct gtc ttc ccc ttc ctt gcc att ggg ggg tac aaa Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Ile Gly Gly Tyr Lys 545 550 555 560			1680
gga aag gac tat tct gag gca gag gcc ctg atc atg acg ttc tcc ctc Gly Lys Asp Tyr Ser Glu Ala Glu Ala Leu Ile Met Thr Phe Ser Leu 565 570 575			1728
aac aat tac cct gcc ggg gac ccc cgt ctg gcc cag gcc aag ctg tgg Asn Asn Tyr Pro Ala Gly Asp Pro Arg Leu Ala Gln Ala Lys Leu Trp 580 585 590			1776
gag gag gcc ttc tta gag gaa atg cga gcc ttc cag cgt cgg atg gct Glu Glu Ala Phe Leu Glu Glu Met Arg Ala Phe Gln Arg Arg Met Ala 595 600 605			1824
ggc atg ttc cag gtc acg ttc acg gct gag cgc tct ctg gaa gac gag Gly Met Phe Gln Val Thr Phe Thr Ala Glu Arg Ser Leu Glu Asp Glu 610 615 620			1872
atc aat cgc acc aca gct gaa gac ctg ccc atc ttt gcc acc agc tac Ile Asn Arg Thr Thr Ala Glu Asp Leu Pro Ile Phe Ala Thr Ser Tyr 625 630 635 640			1920
att gtc ata ttc ctg tac atc tct ctg gcc ctg ggc agc tat tcc agc Ile Val Ile Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Ser 645 650 655			1968
tgg agc cga gtg atg gtg gac tcc aag gcc acg ctg ggc ctc ggc ggg Trp Ser Arg Val Met Val Asp Ser Lys Ala Thr Leu Gly Leu Gly Gly 660 665 670			2016
gtg gcc gtg gtc ctg gga gca gtc atg gct gcc atg ggc ttc ttc tcc Val Ala Val Val Leu Gly Ala Val Met Ala Ala Met Gly Phe Phe Ser 675 680 685			2064
tac ttg ggt atc cgc tcc tcc ctg gtc atc ctg caa gtg gtt cct ttc Tyr Leu Gly Ile Arg Ser Ser Leu Val Ile Leu Gln Val Val Pro Phe 690 695 700			2112
ctg gtg ctg tcc gtg ggg gct gat aac atc ttc atc ttt gtt ctc gag Leu Val Leu Ser Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu 705 710 715 720			2160
tac cag agg ctg ccc cgg agg cct ggg gag cca cga gag gtc cac att			2208

Tyr	Gln	Arg	Leu	Pro	Arg	Arg	Pro	Gly	Glu	Pro	Arg	Glu	Val	His	Ile		
				725					730					735			
ggg	cga	gcc	cta	ggc	agg	gtg	gct	ccc	agc	atg	ctg	ttg	tgc	agc	ctc		2256
Gly	Arg	Ala	Leu	Gly	Arg	Val	Ala	Pro	Ser	Met	Leu	Leu	Cys	Ser	Leu		
			740					745					750				
tct	gag	gcc	atc	tgc	ttc	ttc	cta	ggg	gcc	ctg	acc	ccc	atg	cca	gct		2304
Ser	Glu	Ala	Ile	Cys	Phe	Phe	Leu	Gly	Ala	Leu	Thr	Pro	Met	Pro	Ala		
			755				760					765					
gtg	cgg	acc	ttt	gcc	ctg	acc	tct	ggc	ctt	gca	gtg	atc	ctt	gac	ttc		2352
Val	Arg	Thr	Phe	Ala	Leu	Thr	Ser	Gly	Leu	Ala	Val	Ile	Leu	Asp	Phe		
			770				775					780					
ctc	ctg	cag	atg	tca	gcc	ttt	gtg	gcc	ctg	ctc	tcc	ctg	gac	agc	aag		2400
Leu	Leu	Gln	Met	Ser	Ala	Phe	Val	Ala	Leu	Leu	Ser	Leu	Asp	Ser	Lys		
			785			790				795					800		
agg	cag	gag	gcc	tcc	cgg	ttg	gac	gtc	tgc	tgc	tgt	gtc	aag	ccc	cag		2448
Arg	Gln	Glu	Ala	Ser	Arg	Leu	Asp	Val	Cys	Cys	Cys	Val	Lys	Pro	Gln		
				805					810					815			
gag	ctg	ccc	ccg	cct	ggc	cag	gga	gag	ggg	ctc	ctg	ctt	ggc	ttc	ttc		2496
Glu	Leu	Pro	Pro	Pro	Gly	Gln	Gly	Glu	Gly	Leu	Leu	Leu	Gly	Phe	Phe		
			820					825					830				
caa	aag	gct	tat	gcc	ccc	ttc	ctg	ctg	cac	tgg	atc	act	cga	ggt	gtt		2544
Gln	Lys	Ala	Tyr	Ala	Pro	Phe	Leu	Leu	His	Trp	Ile	Thr	Arg	Gly	Val		
			835				840					845					
gtg	ctg	ctg	ctg	ttt	ctc	gcc	ctg	ttc	gga	gtg	agc	ctc	tac	tcc	atg		2592
Val	Leu	Leu	Leu	Phe	Leu	Ala	Leu	Phe	Gly	Val	Ser	Leu	Tyr	Ser	Met		
			850			855					860						
tgc	cac	atc	agc	gtg	gga	ctg	gac	cag	gag	ctg	gcc	ctg	ccc	aag	gac		2640
Cys	His	Ile	Ser	Val	Gly	Leu	Asp	Gln	Glu	Leu	Ala	Leu	Pro	Lys	Asp		
					870					875					880		
tcg	tac	ctg	ctt	gac	tat	ttc	ctc	ttt	ctg	aac	cgc	tac	ttc	gag	gtg		2688
Ser	Tyr	Leu	Leu	Asp	Tyr	Phe	Leu	Phe	Leu	Asn	Arg	Tyr	Phe	Glu	Val		
				885					890					895			
ggg	gcc	ccg	gtg	tac	ttt	gtt	acc	acc	ttg	ggc	tac	aac	ttc	tcc	agc		2736
Gly	Ala	Pro	Val	Tyr	Phe	Val	Thr	Thr	Leu	Gly	Tyr	Asn	Phe	Ser	Ser		
			900					905					910				
gag	gct	ggg	atg	aat	gcc	atc	tgc	tcc	agt	gca	ggc	tgc	aac	aac	ttc		2784
Glu	Ala	Gly	Met	Asn	Ala	Ile	Cys	Ser	Ser	Ala	Gly	Cys	Asn	Asn	Phe		
			915				920					925					
tcc	ttc	acc	cag	aag	atc	cag	tat	gcc	aca	gag	ttc	cct	gag	cag	tct		2832
Ser	Phe	Thr	Gln	Lys	Ile	Gln	Tyr	Ala	Thr	Glu	Phe	Pro	Glu	Gln	Ser		
			930			935					940						
tac	ctg	gcc	atc	cct	gcc	tcc	tcc	tgg	gtg	gat	gac	ttc	att	gac	tgg		2880
Tyr	Leu	Ala	Ile	Pro	Ala	Ser	Ser	Trp	Val	Asp	Asp	Phe	Ile	Asp	Trp		
					950					955					960		

ctg acc ccg tcc tcc tgc tgc cgc ctt tat ata tct ggc ccc aat aag Leu Thr Pro Ser Ser Cys Cys Arg Leu Tyr Ile Ser Gly Pro Asn Lys 965 970 975	2928
gac aag ttc tgc ccc tcg acc gtc aac tct ctg aac tgc cta aag aac Asp Lys Phe Cys Pro Ser Thr Val Asn Ser Leu Asn Cys Leu Lys Asn 980 985 990	2976
tgc atg agc atc acg atg ggc tct gtg agg ccc tcg gtg gag cag ttc Cys Met Ser Ile Thr Met Gly Ser Val Arg Pro Ser Val Glu Gln Phe 995 1000 1005	3024
cat aag tat ctt ccc tgg ttc ctg aac gac cgg ccc aac atc aaa His Lys Tyr Leu Pro Trp Phe Leu Asn Asp Arg Pro Asn Ile Lys 1010 1015 1020	3069
tgt ccc aaa ggc ggc ctg gca gca tac agc acc tct gtg aac ttg Cys Pro Lys Gly Gly Leu Ala Ala Tyr Ser Thr Ser Val Asn Leu 1025 1030 1035	3114
act tca gat ggc cag gtt tta gcc tcc agg ttc atg gcc tat cac Thr Ser Asp Gly Gln Val Leu Ala Ser Arg Phe Met Ala Tyr His 1040 1045 1050	3159
aag ccc ctg aaa aac tca cag gat tac aca gaa gct ctg cgg gca Lys Pro Leu Lys Asn Ser Gln Asp Tyr Thr Glu Ala Leu Arg Ala 1055 1060 1065	3204
gct cga gag ctg gca gcc aac atc act gct gac ctg cgg aaa gtg Ala Arg Glu Leu Ala Ala Asn Ile Thr Ala Asp Leu Arg Lys Val 1070 1075 1080	3249
cct gga aca gac ccg gct ttt gag gtc ttc ccc tac acg atc acc Pro Gly Thr Asp Pro Ala Phe Glu Val Phe Pro Tyr Thr Ile Thr 1085 1090 1095	3294
aat gtg ttt tat gag cag tac ctg acc atc ctc cct gag ggg ctc Asn Val Phe Tyr Glu Gln Tyr Leu Thr Ile Leu Pro Glu Gly Leu 1100 1105 1110	3339
ttc atg ctc agc ctc tgc ctt gtg ccc acc ttc gct gtc tcc tgc Phe Met Leu Ser Leu Cys Leu Val Pro Thr Phe Ala Val Ser Cys 1115 1120 1125	3384
ctc ctg ctg ggc ctg gac ctg cgc tcc ggc ctc ctc aac ctg ctc Leu Leu Leu Gly Leu Asp Leu Arg Ser Gly Leu Leu Asn Leu Leu 1130 1135 1140	3429
tcc att gtc atg atc ctc gtg gac act gtc ggc ttc atg gcc ctg Ser Ile Val Met Ile Leu Val Asp Thr Val Gly Phe Met Ala Leu 1145 1150 1155	3474
tgg gac atc agt tac aat gct gtg tcc ctc atc aac ctg gtc tcg Trp Asp Ile Ser Tyr Asn Ala Val Ser Leu Ile Asn Leu Val Ser 1160 1165 1170	3519
gcg gtg ggc atg tct gtg gag ttt gtg tcc cac att acc cgc tcc Ala Val Gly Met Ser Val Glu Phe Val Ser His Ile Thr Arg Ser 1175 1180 1185	3564

ttt gcc atc agc acc aag ccc acc tgg ctg gag agg gcc aaa gag 3609
Phe Ala Ile Ser Thr Lys Pro Thr Trp Leu Glu Arg Ala Lys Glu
1190 1195 1200

gcc acc atc tct atg gga agt gcg gtg ttt gca ggt gtg gcc atg 3654
Ala Thr Ile Ser Met Gly Ser Ala Val Phe Ala Gly Val Ala Met
1205 1210 1215

acc aac ctg cct ggc atc ctt gtc ctg ggc ctc gcc aag gcc cag 3699
Thr Asn Leu Pro Gly Ile Leu Val Leu Gly Leu Ala Lys Ala Gln
1220 1225 1230

ctc att cag atc ttc ttc ttc cgc ctc aac ctc ctg atc act ctg 3744
Leu Ile Gln Ile Phe Phe Phe Arg Leu Asn Leu Leu Ile Thr Leu
1235 1240 1245

ctg ggc ctg ctg cat ggc ttg gtc ttc ctg ccc gtc atc ctc agc 3789
Leu Gly Leu Leu His Gly Leu Val Phe Leu Pro Val Ile Leu Ser
1250 1255 1260

tac gtg ggg cct gac gtt aac ccg gct ctg gca ctg gag cag aag 3834
Tyr Val Gly Pro Asp Val Asn Pro Ala Leu Ala Leu Glu Gln Lys
1265 1270 1275

cgg gct gag gag gcg gtg gca gca gtc atg gtg gcc tct tgc cca 3879
Arg Ala Glu Glu Ala Val Ala Ala Val Met Val Ala Ser Cys Pro
1280 1285 1290

aat cac ccc tcc cga gtc tcc aca gct gac aac atc tat gtc aac 3924
Asn His Pro Ser Arg Val Ser Thr Ala Asp Asn Ile Tyr Val Asn
1295 1300 1305

cac agc ttt gaa ggt tct atc aaa ggt gct ggt gcc atc agc aac 3969
His Ser Phe Glu Gly Ser Ile Lys Gly Ala Gly Ala Ile Ser Asn
1310 1315 1320

ttc ttg ccc aac aat ggg cgg cag ttc tga 3999
Phe Leu Pro Asn Asn Gly Arg Gln Phe
1325 1330

<210> 4

<211> 1332

<212> PRT

<213> Homo sapiens

<400> 4

Met Ala Glu Ala Gly Leu Arg Gly Trp Leu Leu Trp Ala Leu Leu Leu
1 5 10 15

Arg Leu Ala Gln Ser Glu Pro Tyr Thr Thr Ile His Gln Pro Gly Tyr

20	25	30
Cys Ala Phe Tyr Asp Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Ser		
35	40	45
Leu Met Thr Leu Ser Asn Val Ser Cys Leu Ser Asn Thr Pro Ala Arg		
50	55	60
Lys Ile Thr Gly Asp His Leu Ile Leu Leu Gln Lys Ile Cys Pro Arg		
65	70	75
Leu Tyr Thr Gly Pro Asn Thr Gln Ala Cys Cys Ser Ala Lys Gln Leu		
85	90	95
Val Ser Leu Glu Ala Ser Leu Ser Ile Thr Lys Ala Leu Leu Thr Arg		
100	105	110
Cys Pro Ala Cys Ser Asp Asn Phe Val Asn Leu His Cys His Asn Thr		
115	120	125
Cys Ser Pro Asn Gln Ser Leu Phe Ile Asn Val Thr Arg Val Ala Gln		
130	135	140
Leu Gly Ala Gly Gln Leu Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr		
145	150	155
Gln His Ser Phe Ala Glu Gln Ser Tyr Asp Ser Cys Ser Arg Val Arg		
165	170	175
Val Pro Ala Ala Ala Thr Leu Ala Val Gly Thr Met Cys Gly Val Tyr		
180	185	190
Gly Ser Ala Leu Cys Asn Ala Gln Arg Trp Leu Asn Phe Gln Gly Asp		
195	200	205
Thr Gly Asn Gly Leu Ala Pro Leu Asp Ile Thr Phe His Leu Leu Glu		
210	215	220
Pro Gly Gln Ala Val Gly Ser Gly Ile Gln Pro Leu Asn Glu Gly Val		
225	230	235
Ala Arg Cys Asn Glu Ser Gln Gly Asp Asp Val Ala Thr Cys Ser Cys		
245	250	255

Gln Asp Cys Ala Ala Ser Cys Pro Ala Ile Ala Arg Pro Gln Ala Leu
260 265 270

Asp Ser Thr Phe Tyr Leu Gly Gln Met Pro Gly Ser Leu Val Leu Ile
275 280 285

Ile Ile Leu Cys Ser Val Phe Ala Val Val Thr Ile Leu Leu Val Gly
290 295 300

Phe Arg Val Ala Pro Ala Arg Asp Lys Ser Lys Met Val Asp Pro Lys
305 310 315 320

Lys Gly Thr Ser Leu Ser Asp Lys Leu Ser Phe Ser Thr His Thr Leu
325 330 335

Leu Gly Gln Phe Phe Gln Gly Trp Gly Thr Trp Val Ala Ser Trp Pro
340 345 350

Leu Thr Ile Leu Val Leu Ser Val Ile Pro Val Val Ala Leu Ala Ala
355 360 365

Gly Leu Val Phe Thr Glu Leu Thr Thr Asp Pro Val Glu Leu Trp Ser
370 375 380

Ala Pro Asn Ser Gln Ala Arg Ser Glu Lys Ala Phe His Asp Gln His
385 390 395 400

Phe Gly Pro Phe Phe Arg Thr Asn Gln Val Ile Leu Thr Ala Pro Asn
405 410 415

Arg Ser Ser Tyr Arg Tyr Asp Ser Leu Leu Leu Gly Pro Lys Asn Phe
420 425 430

Ser Gly Ile Leu Asp Leu Asp Leu Leu Leu Glu Leu Leu Glu Leu Gln
435 440 445

Glu Arg Leu Arg His Leu Gln Val Trp Ser Pro Glu Ala Gln Arg Asn
450 455 460

Ile Ser Leu Gln Asp Ile Cys Tyr Ala Pro Leu Asn Pro Asp Asn Thr
465 470 475 480

Ser Leu Tyr Asp Cys Cys Ile Asn Ser Leu Leu Gln Tyr Phe Gln Asn
485 490 495

Asn Arg Thr Leu Leu Leu Leu Thr Ala Asn Gln Thr Leu Met Gly Gln
500 505 510

Thr Ser Gln Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala
515 520 525

Pro Leu Thr Phe Lys Asp Gly Thr Ala Leu Ala Leu Ser Cys Met Ala
530 535 540

Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Ile Gly Gly Tyr Lys
545 550 555 560

Gly Lys Asp Tyr Ser Glu Ala Glu Ala Leu Ile Met Thr Phe Ser Leu
565 570 575

Asn Asn Tyr Pro Ala Gly Asp Pro Arg Leu Ala Gln Ala Lys Leu Trp
580 585 590

Glu Glu Ala Phe Leu Glu Glu Met Arg Ala Phe Gln Arg Arg Met Ala
595 600 605

Gly Met Phe Gln Val Thr Phe Thr Ala Glu Arg Ser Leu Glu Asp Glu
610 615 620

Ile Asn Arg Thr Thr Ala Glu Asp Leu Pro Ile Phe Ala Thr Ser Tyr
625 630 635 640

Ile Val Ile Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Ser
645 650 655

Trp Ser Arg Val Met Val Asp Ser Lys Ala Thr Leu Gly Leu Gly Gly
660 665 670

Val Ala Val Val Leu Gly Ala Val Met Ala Ala Met Gly Phe Phe Ser
675 680 685

Tyr Leu Gly Ile Arg Ser Ser Leu Val Ile Leu Gln Val Val Pro Phe
690 695 700

Leu Val Leu Ser Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu
705 710 715 720

Tyr Gln Arg Leu Pro Arg Arg Pro Gly Glu Pro Arg Glu Val His Ile
725 730 735

Gly Arg Ala Leu Gly Arg Val Ala Pro Ser Met Leu Leu Cys Ser Leu
740 745 750

Ser Glu Ala Ile Cys Phe Phe Leu Gly Ala Leu Thr Pro Met Pro Ala
755 760 765

Val Arg Thr Phe Ala Leu Thr Ser Gly Leu Ala Val Ile Leu Asp Phe
770 775 780

Leu Leu Gln Met Ser Ala Phe Val Ala Leu Leu Ser Leu Asp Ser Lys
785 790 795 800

Arg Gln Glu Ala Ser Arg Leu Asp Val Cys Cys Cys Val Lys Pro Gln
805 810 815

Glu Leu Pro Pro Pro Gly Gln Gly Glu Gly Leu Leu Leu Gly Phe Phe
820 825 830

Gln Lys Ala Tyr Ala Pro Phe Leu Leu His Trp Ile Thr Arg Gly Val
835 840 845

Val Leu Leu Leu Phe Leu Ala Leu Phe Gly Val Ser Leu Tyr Ser Met
850 855 860

Cys His Ile Ser Val Gly Leu Asp Gln Glu Leu Ala Leu Pro Lys Asp
865 870 875 880

Ser Tyr Leu Leu Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Phe Glu Val
885 890 895

Gly Ala Pro Val Tyr Phe Val Thr Thr Leu Gly Tyr Asn Phe Ser Ser
900 905 910

Glu Ala Gly Met Asn Ala Ile Cys Ser Ser Ala Gly Cys Asn Asn Phe
915 920 925

Ser Phe Thr Gln Lys Ile Gln Tyr Ala Thr Glu Phe Pro Glu Gln Ser
930 935 940

Tyr Leu Ala Ile Pro Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp
945 950 955 960

Leu Thr Pro Ser Ser Cys Cys Arg Leu Tyr Ile Ser Gly Pro Asn Lys

965										970					975				
Asp	Lys	Phe	Cys	Pro	Ser	Thr	Val	Asn	Ser	Leu	Asn	Cys	Leu	Lys	Asn				
			980					985					990						
Cys	Met	Ser	Ile	Thr	Met	Gly	Ser	Val	Arg	Pro	Ser	Val	Glu	Gln	Phe				
		995					1000					1005							
His	Lys	Tyr	Leu	Pro	Trp	Phe	Leu	Asn	Asp	Arg	Pro	Asn	Ile	Lys					
	1010					1015					1020								
Cys	Pro	Lys	Gly	Gly	Leu	Ala	Ala	Tyr	Ser	Thr	Ser	Val	Asn	Leu					
	1025					1030					1035								
Thr	Ser	Asp	Gly	Gln	Val	Leu	Ala	Ser	Arg	Phe	Met	Ala	Tyr	His					
	1040					1045					1050								
Lys	Pro	Leu	Lys	Asn	Ser	Gln	Asp	Tyr	Thr	Glu	Ala	Leu	Arg	Ala					
	1055					1060					1065								
Ala	Arg	Glu	Leu	Ala	Ala	Asn	Ile	Thr	Ala	Asp	Leu	Arg	Lys	Val					
	1070					1075					1080								
Pro	Gly	Thr	Asp	Pro	Ala	Phe	Glu	Val	Phe	Pro	Tyr	Thr	Ile	Thr					
	1085					1090					1095								
Asn	Val	Phe	Tyr	Glu	Gln	Tyr	Leu	Thr	Ile	Leu	Pro	Glu	Gly	Leu					
	1100					1105					1110								
Phe	Met	Leu	Ser	Leu	Cys	Leu	Val	Pro	Thr	Phe	Ala	Val	Ser	Cys					
	1115					1120					1125								
Leu	Leu	Leu	Gly	Leu	Asp	Leu	Arg	Ser	Gly	Leu	Leu	Asn	Leu	Leu					
	1130					1135					1140								
Ser	Ile	Val	Met	Ile	Leu	Val	Asp	Thr	Val	Gly	Phe	Met	Ala	Leu					
	1145					1150					1155								
Trp	Asp	Ile	Ser	Tyr	Asn	Ala	Val	Ser	Leu	Ile	Asn	Leu	Val	Ser					
	1160					1165					1170								
Ala	Val	Gly	Met	Ser	Val	Glu	Phe	Val	Ser	His	Ile	Thr	Arg	Ser					
	1175					1180					1185								

Phe Ala Ile Ser Thr Lys Pro Thr Trp Leu Glu Arg Ala Lys Glu
 1190 1195 1200
 Ala Thr Ile Ser Met Gly Ser Ala Val Phe Ala Gly Val Ala Met
 1205 1210 1215
 Thr Asn Leu Pro Gly Ile Leu Val Leu Gly Leu Ala Lys Ala Gln
 1220 1225 1230
 Leu Ile Gln Ile Phe Phe Phe Arg Leu Asn Leu Leu Ile Thr Leu
 1235 1240 1245
 Leu Gly Leu Leu His Gly Leu Val Phe Leu Pro Val Ile Leu Ser
 1250 1255 1260
 Tyr Val Gly Pro Asp Val Asn Pro Ala Leu Ala Leu Glu Gln Lys
 1265 1270 1275
 Arg Ala Glu Glu Ala Val Ala Ala Val Met Val Ala Ser Cys Pro
 1280 1285 1290
 Asn His Pro Ser Arg Val Ser Thr Ala Asp Asn Ile Tyr Val Asn
 1295 1300 1305
 His Ser Phe Glu Gly Ser Ile Lys Gly Ala Gly Ala Ile Ser Asn
 1310 1315 1320
 Phe Leu Pro Asn Asn Gly Arg Gln Phe
 1325 1330

<210> 5

<211> 885

<212> DNA

<213> Rattus sp.

<400> 5

ccacgcgtcc gcacctgcaa gtgtggtccc ctgaggcaga gcgcaacatc tccctccagg 60
 acatctgcta tgccccctc aaccatata acaccagcct ctccgactgc tgtgtcaaca 120
 gcctccttca gtacttccag aacaaccgca cctcctgat gctcacggcc aaccagactc 180
 tgaatggcca gacctccctg gtggactgga aggaccattt cctctactgt gcaaagccc 240

ctctcacgtt caaagatggc acgtctctgg ccctgagctg catggctgac tacggggctc 300
ctgtcttccc cttecttgct gttgggggat accaaggcac ggactattcc gaggcagaag 360
cgctgatcat aaccttctct ctcaataact accccgctga tgatccccgc atggcccagg 420
ccaagctctg ggaggaggct ttcttgaagg aaatggaatc cttccagagg aacacaagtg 480
acaagttcca ggttgcggtc tcagctgagc gctctctgga ggatgagatc aaccgcacca 540
ccatccagga cctgcctgtc tttgccgtca gctacattat cgtcttctctg tacatctccc 600
tggccttggg cagctactcc agatgcagcc gagtagcggg ggagtccaag gctactctgg 660
gcctaggtgg ggtgatagtg tgctgggagc agttctggct tgcattggggc ttctaactcc 720
tacctgggtg tcccctcttc tctgggtatc atccaagtgg tacctttcct ggtgcttaag 780
ctgtgggagc tggacacatc tacatcctag acttgagtac cagaggtacc taggaagccg 840
cggaacagcg aaaaggacac attgggcgca ccctgggcat gtggc 885

<210> 6

<211> 458

<212> DNA

<213> Rattus sp.

<400> 6

gaccagatgt taaccaagct ctggtacagg aggagaaact agccagcagag gcagcagtgg 60
ccccagagcc ttcttgccca cagtaccctt cccctgctga tgcggatgcc aatgttaact 120
acggctttgc cccagaactt gccacaggag ctaatgctgc tagaagctct ttgcccaaaa 180
gtgacaaaaa gttctaattg agtaggagct tgtccatgct tctgctgatg agggatcatg 240
aaggtcttcc ctctggttgt cctcaaggcc tggggggagg ttgttcagag aaaaatggct 300
ggcattcctg ccaagaggca accggcagct tggcactgac tccttgggtct cataggtccc 360
taaggcttgg tcagattact cctcatggag agactatctt aagtatctaa gctatcgatt 420
gggatgcacg gctgttcatt aaaaaggcta tggctatg 458

<210> 7

<211> 896

<212> DNA

<213> Rattus sp.

<400> 7

ccacgcgtcc gcagtttcat aagtacctgc cctgggttct gaatgatccg cccaatatca	60
gatgtcccaa aggggggtcta gcagcgtata gaacgtctgt gaatttgagc tcagatggcc	120
aggttatagc ctcccagttc atggcctacc acaagccctt aagggaactca caggacttca	180
cagaagctct ccgggcgtcc cggttgctag cagccaacat cacagctgac ctacggaagg	240
tgcctgggac agatccaaac tttgaggtct tcccttacac gatctccaac gtgttctacc	300
agcaatacct gacggtcctt cctgagggaa tcttcaccct tgcctcttgc tttgtgcca	360
cctttgttgt ctgctacctc ctactgggoc tggacatgtg ctacgggatc ctcaacctac	420
tctccatcat tatgattctc gtggacacca ttggcctcat ggctgtgtgg ggtatcagct	480
ataatgcggg atccctcatc aaccttgtca cggcagtgagg catgtctgtg gagtttgtgt	540
cccacatcac tcggtccttt gcttgtaagc accaagccta cccggctgga gagggctaaa	600
agatgctact gtcttcatgg gcagtgcggg gtttgctgga gtggccatga ccaacttccc	660
aggcatcctc atcttggggg ctttgcccca agccagggc tattcagatc ttcttcttcc	720
gcctcaacct tctgatcacc tttgctgggg tctgctgcat ggctggctct cctgcccggg	780
ttgtcctcag ctatctggga ccagatgtaa ccaaggctct gctaccggga ggagaaacta	840
gccagcgagg gcagcagtgg cccagagac ttcttgccca caagtaccct tccctg	896

<210> 8

<211> 3124

<212> DNA

<213> Rattus sp.

<400> 8

tgcaagtgtg gtcccctgag gcagagcgca acatctccct ccaggacatc tgctatgccc	60
ccctcaaccc atataacacc agcctctccg actgctgtgt caacagcctc cttcagtact	120
tccagaacaa ccgcaccctc ctgatgctca cggccaacca gactctgaat ggccagacct	180
ccctgggtgga ctggaaggac catttcctct actgtgcaaa tgcccctctc acgttcaaag	240
atggcacgtc tctggccctg agctgcatgg ctgactacgg ggctcctgtc ttccccttcc	300
ttgctgttgg gggataccaa ggcacggact attccgaggc agaagcgtg atcataacct	360
tctctctcaa taactacccc gctgatgatc cccgcatggc ccaggccaag ctctgggagg	420

aggctttctt gaaggaaatg gaatccttcc agaggaacac aagtgacaag ttccagggtg 480
cggttctcagc tgagcgctct ctggaggatg agatcaaccg caccaccatc caggacctgc 540
ctgtctttgc cgtcagctac attatcgtct tcctgtacat ctccctggcc ctgggcagct 600
actccagatg cagccgagta gcggtggagt ccaaggctac tctgggecta ggtgggggtga 660
ttgttggtgt gggagcagtt ctggctgcca tgggcttcta ctectacctg ggtgtcccct 720
cttctctggt tatcatccaa gtggtacctt tcctggtgct agctgtggga gctgacaaca 780
tcttcatctt tgttcttgag taccagaggc tacctaggat gcctggggaa cagcgagagg 840
ctcacattgg ccgcacctg ggcagtgtgg cccccagcat gctgctgtgc agcctctctg 900
aggccatctg cttctttcta ggggccctga ccccatgcc agctgtgagg accttcgcct 960
tgacctctgg cttagcaatt atcctcgact tcctgctcca gatgactgcc tttgtggccc 1020
tgctctccct ggatagcaag aggcaggagg cctctcgccc ggatgtctta tgctgctttt 1080
caaccoggaa gctgccccca cctaaagaaa aagaaggcct ctactccgc ttcttccgca 1140
agatatacgc tcctttcctg ctgcacagat tcatccgccc tgttgatgatg ctgctgtttc 1200
tgacctgtt tggagcaaat ctctacttaa tgtgcaacat caacgtgggg ctagaccagg 1260
agctggctct gcccaaggac tcgtacttga tagactactt cctctttctg aaccgatacc 1320
ttgaagtggg gcctccagtg tactttgtca ccacctggg cttcaacttc tccagcgagg 1380
caggcatgaa cgccacttgc tctagcgcag gctgtaagag cttctcccta acccagaaaa 1440
tccagtatgc cagtgaattc cctgaccagt cttacgtggc tattgctgca tcctcctggg 1500
tagatgactt catcgactgg ctgaccccg cctcctcctg ctgtcgcctt tatatacgtg 1560
gccccataa ggatgagttc tgtccctcaa cggatacttc cttcaactgc ttaaaaaact 1620
gcatgaaccg cactctgggt cctgtgaggc ccacagcgga acagtttcat aagtacctgc 1680
cctggttccct gaatgatccg cccaatatca gatgtcccaa agggggctta gcagcgtata 1740
gaacgtctgt gaatttgagc tcagatggcc aggttatagc ctcccagttc atggcctacc 1800
acaagccctt aaggaactca caggacttca cagaagctct ccgggcgtcc cggttgctag 1860
cagccaacat cacagctgac ctacggaagg tgctgggac agatccaaac tttgaggtct 1920
tcccttacac gatctccaac gtgttctacc agcaatacct gacggtcctt cctgagggaa 1980
tcttcaccct tgctctttgc tttgtgcca ctttgttgt ctgtacctc ctactgggcc 2040
tggacatgtg ctcagggatc ctcaacctac tctccatcat tatgattctc gtggacacca 2100
ttggcctcat ggctgtgtgg ggtatcagct ataatgcggt atccctcacc aaccttgtca 2160
cggcagtggg catgtctgtg gagtttgtgt ccacatcac tcggtccttt gctgtaagca 2220

ccaagcctac ccggctggag agggctaaag atgctactgt cttcatgggc agtgcggtgt 2280
ttgctggagt ggccatgacc aacttcccag gcattctcat cttgggcttt gcccaagccc 2340
agcttattca gatcttcttc ttccgcctca accttctgat caccttgctg ggtctgctgc 2400
atggcctggg cttcctgccg gttgtcctca gctatctggg accagatgtt aaccaagctc 2460
tggtacagga ggagaaacta gccagcgagg cagcagtggc ccagagcct tcttgccac 2520
agtaccctc ccctgctgat gcggatgcca atgttaacta cggttttgc ccagaacttg 2580
cccacggagc taatgctgct agaagctctt tgcccaaaag tgaccaaaag ttctaattga 2640
gtaggagctt gtccatgctt cttgctgatg agggatcatg aaggtcttcc ctctggttgt 2700
cctcaaggcc tggggggagg ttgtttcaga gaaaaatggc tggcattcct gccacgaggc 2760
aaccggcagc attggcactg acctccttgc tctcataggt ccctaaggcc ttggtcagat 2820
tacctcctcc atggagagac tatcttaagt atcttaagta tcgtatggga tgcctgcct 2880
gtcaattaaa aaggctatgg cctatggctc aggcagggcc atccggaaga agagaggatt 2940
ctgggataaa gccagggtggg agattcgcct ggggaaaatg tgacaatggg tcctgagcat 3000
gggcaatcag ccatgtggca gaatgtaaat taatataaat gggttgtctt aagttatgat 3060
tctagctggg gaggagccta gctgtgtage caagatattt gtaaatataa aaaaaaaaaa 3120
aaaa 3124

<210> 9

<211> 4484

<212> DNA

<213> Rattus sp.

<400> 9

atggcagctg cctggctggg atggctgctc tgggccctgc tcctgagcgc ggcccagggt 60
gagctataca caccacaaca cgaagctggg gtctgcacct ttacgaaga gtgcgggaaa 120
aaccagagc tctctggagg cctcacgtca ctatccaatg tatcctgcct gtctaacc 180
ccggcccgcc acgtcacggg tgaacacctg gctcttctcc agcgcattctg tccccgcctg 240
tacaacggcc ccaataccac ttttgctgtg tgctctacca agcagctgct gtccttagaa 300
agcagcatgt ccatcaccaa ggcccttctc acgcgctgcc cggcctgctc tgacaatttt 360
gtgagcttac actgccaca cacttgagc cctgaccaga gcctcttcat caacgtcacc 420

cggttggttg agcggggcgc tggagagcct cctgccgtgg tggcctatga ggccttttat 480
cagcgcagct ttgctgagaa ggcctatgag tcctgcagcc aggtgcgcat ccctgcggcc 540
gcttccttgg ccgtgggcag catgtgtgga gtgtatggct ccgccctctg caatgctcag 600
cgctggctca acttccaagg agacacaggg aatggcctgg ctccgctgga taccaccttc 660
cacctcttgg agcctggcca ggcctaccg gatgggatcc agccactgaa tgggaagatc 720
gcaccctgca acgagtctca gggatgatgac tcagcagctc gctcctgcca ggactgtgcg 780
gcgtcctgcc ctgtcatccc tccgcccagag gccttgccgc cttccttcta catgggtcgc 840
atgccaggct ggctggccct catcatcatc ttcactgctg tctttgtgtt gctctctgca 900
gtccttgtgc gtctccgagt ggtttccaac aggaacaaga acaaggcaga agggccccag 960
gaagccccca aactccctca taagcacaaa ctctcacccc ataccatcct gggccggttc 1020
ttccagaact ggggcacaag ggtggcctcg tggccactca ccgtcttagc actgtccttc 1080
atcgttgtga tagccttagc agcaggcctg acctttattg aactcaccac agaccctgtg 1140
gaactgtggt cggcccccaa gagccaggcc cggaaagaga agtctttcca tgatgagcat 1200
ttcggccctc tctttcgaac caaccagatt ttcgtgacag ctccgaacag gtccagctac 1260
aagtacgact ccctactgct agggccaag aacttcagtg ggatcctgtc cctggacttc 1320
ctgctggagc tgctggagct tcaggagagg ctccgacacc tgcaagtgtg gtccctgag 1380
gcagagcgca acatctccct ccaggacatc tgctatgccc ccctcaacce atataacacc 1440
agcctctccg actgtgtgtt caacagcctc cttcagtact tccagaacaa ccgcaccctc 1500
ctgatgctca cggccaacca gactctgaat ggccagacct ccctgggtgga ctggaaggac 1560
catttcctct actgtgcaaa tgcccctctc acgttcaaag atggcacgtc tctggccctg 1620
agctgcatgg ctgactacgg ggctcctgtc tcccccttcc ttgctgttgg gggataccaa 1680
ggcacggact attccgaggc agaagcgctg atcataacct tctctctcaa taactacccc 1740
gctgatgatc cccgcatggc ccaggccaag ctctgggagg aggccttctt gaaggaaatg 1800
gaatccttcc agaggaacac aagtgacaag ttccagggtt cgcttctcagc tgagcgctct 1860
ctggaggatg agatcaaccg caccaccatc caggacctgc ctgtctttgc cgtcagctac 1920
attatcgtct tcctgtacat ctccctggcc ctgggcagct actccagatg cagccgagta 1980
gcggtggagt ccaaggctac tctgggccta ggtggggtga ttgttgtgct gggagcagtt 2040
ctggctgcca tgggtctcta ctctacctg ggtgtccct cttctctggt taccatccaa 2100
gtggtacctt tcctggtgct agctgtggga gctgacaaca tcttcactct tgttcttgag 2160
taccagaggc taccatgagat gcctggggaa cagcgagagg ctacattgg ccgcaccctg 2220

ggcagtgtgg cccccagcat gctgctgtgc agcctctctg aggccatctg cttcttttcta 2280
ggggccctga ccccatgcc agctgtgagg accttcgcct tgacctctgg cttagcaatt 2340
atcctcgact tcctgctcca gatgactgcc tttgtggccc tgctctccct ggatagcaag 2400
aggcaggagg cctctcgccc ggatgtctta tgctgctttt caaccggaa gctgccccca 2460
cctaaagaaa aagaaggcct cttactccgc ttcttccgca agatatacgc tcctttcctg 2520
ctgcacagat tcatccgccc tgttgtgatg ctgctgtttc tgacctgtt tggagcaaat 2580
ctctacttaa tgtgcaacat caacgtgggg ctagaccagg agctggctct gcccaaggac 2640
tcgtacttga tagactactt cctctttctg aaccgatacc ttgaagtggg gcctccagtg 2700
tactttgtca ccacctcggg cttcaacttc tccagcgagg caggcatgaa cgccacttgc 2760
tctagcgcag gctgtaagag cttctcccta acccagaaaa tccagtatgc cagtgaattc 2820
cctgaccagt cttabgtggc tattgctgca tcctcctggg tagatgactt catcgactgg 2880
ctgaccccggt cctcctcctg ctgtcgcctt tatatacgtg gccccataa ggatgagttc 2940
tgtccctcaa cggatacttc cttcaactgc ttaaaaaact gcatgaaccg cactctgggt 3000
cctgtgaggg ccacagcgga acagtttcat aagtacctgc cctggttcct gaatgatccg 3060
cccaatatca gatgtcccaa aggggggtcta gcagcgtata gaacgtctgt gaatttgagc 3120
tcagatggcc aggttatagc ctcccagttc atggcctacc acaagccctt aaggaactca 3180
caggacttca cagaagctct ccgggcgtcc cggttgctag cagccaacat cacagctgac 3240
ctacggaagg tgccctgggac agatccaaac tttgaggtct tcccttacac gatctccaac 3300
gtgttctacc agcaatacct gacggtcctt cctgagggaa tcttcaccct tgctctttgc 3360
tttgtgccc cttttgttgt ctgctacctc ctactgggcc tggacatgtg ctcagggatc 3420
ctcaacctac tctccatcat tatgattctc gtggacacca ttggcctcat ggctgtgtgg 3480
ggatatcagct ataatgogg atccctcatc aaccttgtca cggcagtggg catgtctgtg 3540
gagtttgtgt ccacatcac tcggtccttt gctgtaagca ccaagcctac ccggctggag 3600
agggctaaag atgctactgt cttcatgggc agtgcggtgt ttgctggagt ggccatgacc 3660
aacttcccag gcctcctcat cttgggcttt gcccaagccc agcttattca gatcttcttc 3720
ttccgcctca accttctgat caccttgctg ggtctgctgc atggcctggg cttcctgccg 3780
gttgtcctca gctatctggg accagatggt aaccaagctc tggtagagga ggagaaacta 3840
gccagcgagg cagcagtggc ccagagcct tcttgccac agtaccctc ccctgctgat 3900
gcggatgcca atgttaacta cggctttgcc ccagaacttg ccacggagc taatgctgct 3960

agaagctctt tgcccaaaag tgaccaaaag ttctaattgga gtaggagctt gtccatgctt 4020
cttgctgatg agggatcatg aaggctctcc ctctggttgt cctcaaggcc tggggggagg 4080
ttgtttcaga gaaaaatggc tggcattcct gccacgaggc aaccggcagc attggcactg 4140
acctccttgc tctcataggt ccctaaggcc ttggtcagat tacctcctcc atggagagac 4200
tatcttaagt atcttaagta tcgtatggga tgcacgcct gtcaattaaa aaggctatgg 4260
cctatggctc aggcagggcc atccggaaga agagaggatt ctgggataaa gccaggtggg 4320
agattcgcct ggggaaaatg tgacaatggg tcctgagcat gggcaatcag ccatgtggca 4380
gaatgtaaataa taatataaat gggttgtctt aagttatgat tctagctggg gaggagccta 4440
gctgtgtagc caagatatatt gtaaatataa aaaaaaaaaa aaaa 4484

<210> 10

<211> 3993

<212> DNA

<213> Rattus sp.

<220>

<221> misc_feature

<222> (1)..(3993)

<223> n is g or a or t or c

<400> 10

atggcngcng cntggytngg ntggyt nytn tgggcnytny tnytnwsngc ngcncarggn 60
garyntaya cncnaarca ygargcnggn gtntgyacnt tytaygarga rtgyggnaar 120
aayccngary tnwsngngg nytnacnwsn ytnwsnaayg tnwsntgyyt nwsnaayacn 180
ccngcnmgnc aygtnacngg ngarcayytn gcnynytnc armgnathtg yccnmgnytn 240
tayaayggnc cnaayacnac nttygcntgy tgywsnacna arcarytnyt nwsnytn gar 300
wsnwsnatgw snathacnaa rgcnyt nytn acnmgtgyc cngcntgyws ngayaaytty 360
gtnwsnytn aytgycayaa yacntgywsn ccngaycarw snytnnttyat haaygtnacn 420
mgngtngtng armngngngc ngngarccn ccngcngtng tngcntayga rgcnttytay 480
carmgnwsnt tygcngaraa rgcntaygar wsntgywsnc argtnmgntat hccngcngcn 540
gcnwsnytn cngtnggnws natgtgyggg gtntayggw sngcnytn tg yaaygcncar 600
mgntggytna ayttycargg ngayacnggn aayggnytn cncnytn ga yathacntty 660
cayytnytn arcngngnaa rgcnytnccn gayggnathe arccnytnaa yggnaarath 720

gcncntgya aygarwsnca rggngaygay wsngcngtnt gywsntgyca rgaytgygc	780
gcnwsntgyc cngtnathcc nccncngar gcnytnmgnc cnwsnttyta yatgggnmgn	840
atgcnggnt ggytngcnyt nathathath ttyacngcng tnttygtnyt nytnwsngcn	900
gtnytngtnm gnytnmgnt ngtnwsnaay mgnaayaara ayaargcnga rggncncar	960
gargcncna arytnccna yaarcayaar ytnwsnccnc ayacnathyt nggnmgntty	1020
ttycaraayt ggggnacnmg ngtnngcnwsn tggcnytna cngtnytngc nytnwsntty	1080
athgtngtna thgcnytnge ngcnggnytn acnttyathg arytnacnac ngayccngtn	1140
garytntggw sngcncnaa rwsncargcn mgnaargara arwsnttyca ygaygarcay	1200
ttyggncnt tyttymgnac naaycarath ttygtnacng cnmgnaaymg nwsnwsntay	1260
aartaygayw snytnytnyt nggnwsnaar aayttywsng gnathytnws nytngaytty	1320
ytnytngary tnytngaryt ncargarmgn ytnmgncayy tncargtntg gwsnccngar	1380
gcngarmgna ayathwsnyt ncargayath tgytaygcnc cnytnaaycc ntayaayacn	1440
wsnytnwsng aytgytgygt naaywsnytn ytncartayt tycaraayaa ymgncnytn	1500
ytnatgytna cngcnaayca racnytnaay ggncaracnw snytngtnga ytggargay	1560
cayttytnt aytgygcnaa ygcncnytn acnttyaarg ayggncnws nytngcnytn	1620
wsntgyatgg cngaytaygg ngcncngtn ttycenttyy tngcngtngg nggntaycar	1680
ggncngayt aywsngargc ngargcnytn athathacnt tywsnytnaa yaaytayccn	1740
gcngaygayc cnmgnatggc ncargcnaar ytntgggarg argcnttyyt naargaratg	1800
garwsnttyc armgnaayac nwsngayaar ttycargtnge cnttywsnge ngarmgnwsn	1860
ytnngargayg arathaaymg nacnacnath cargaytnc cngtnttygc ngtnwsntay	1920
athathgtnt tyytntayat hwsnytnge ytnngnwsnt aywsnmngtg ywsnmngtn	1980
gcngtngarw snaargcnac nytnggnytn ggngngtna thgtngtnyt nggngcngtn	2040
ytngecnaa tggnttyta ywsntayytn ggngtnccnw snwsnytnge nathathcar	2100
gtngtnccnt tyytngtnyt ngcngtnggn gcngayaaya thttiyathht ygtnytngar	2160
taycarmgny tnccnmgnat gccngngar carmngarg cncayathgg nmgnacnytn	2220
ggnwsngtnge cncnwsnat gytnytntgy wsnytnwsng argcnathgt ytttytytn	2280
ggngcnytna cncnatgcc ngcngtnmgn acnttygcny tnacnwsngg nytngcnath	2340
athytngayt tyytnytnca ratgacngcn ttygtngcny tnytnwsnyt ngaywsnaar	2400
mgncargarg cnwsnmgncc ngaygtnytn tgytgyttyw snacnmgnaa rytncncn	2460

ccnaargara argarggnyt nytnytnmgn ttyttymgna arathtaygc nccnttyytn 2520
ytncaymgnt tyathmgnc ngtnngnatg ytntnttyy tnacnytnntt yggngcnaay. 2580
ytntayytna tgtgyaayat haaygtnggn ytngaycarg arytnngcnyt nccnaargay 2640
wsntayytna thgaytaytt yyntnttyytn aaymgntayy tngargtngg nccncngtn 2700
tayttygtna cnacnwsngg nttyaaytty wsnwsngarg cnggnatgaa ygcnaentgy 2760
wsnwsngcng gntgyaarws nttywsnytn acncaraara thcartaygc nwsngartty 2820
ccngaycarw sntaygtngc nathgcngcn wsnwsntggg tngaygaytt yathgaytgg 2880
ytncnccnw snwsnwsntg ytgymgnytn tayathmgng gncncayaa rgaygartty 2940
tgyccnwsna cngayacnws nttyaaytgy ytnaaraayt gyatgaaymg nacnytnngn 3000
cngtnmgnc cnacngcnga rcarttycay aartayytn cngtggttyt naaygayccn 3060
ccnaayathm gntgyccnaa rggnggnytn gcngcntaym gnacnwsngt naayytnwsn 3120
wsngayggnc argtnathgc nwsncartty atggcntayc ayaarccnyt nmgnaywsn 3180
cargayttya cngargcnyt nmngcnwsn mgnytnytn gngcnaayat hacngcngay 3240
ytmgnaarg tncnggnac ngayccnaay ttygargtnt tyccntayac nathwsnaay 3300
gtnttytayc arcartayyt nacngtnytn ccngarggna thttyacnyt ngcnytnngy 3360
ttygtncna cnttygtngt ntgytayytn ytntnggny tngayatgtg ywsnggnath 3420
ytnaayytny tnwsnathat hatgathytn gtngayacna thggnytnat ggngtntgg 3480
ggathwsnt ayaaygcngt nwsnytnath aayytnngtna cngcngtngg natgwsngtn 3540
garttygtnw snayathac nmgnwsntty gcngtnwsna cnaarccnac nmgnytngar 3600
mgngcnaarg aygcnaengt nttyatgggn wsngcngtnt tygcnggngt ngcnatgacn 3660
aayttyccng gnathytnat hytnggntty gcncargcnc arytnathca rathttyty 3720
tymgnytna ayytnytnat hacnytnytn ggnytnytn cnggnytnngt nttyytnccn 3780
gtngtnytnw sntayytnng nccngaygt nccngaycny tngtncarga rgaraarytn 3840
gcnwsngarg cngcngtngc nccngarccn wsntgyccnc artayccnws nccngcngay 3900
gcngaygcna aygtnaayta yggnttygcn ccngarytn cncayggngc naaygcngcn 3960
mgnwsnwsny tncnaarws ngaycaraar tty 3993

<210> 11

<211> 4002

<212> DNA

<213> Mus sp.

<220>

<221> CDS

<222> (1)..(4002)

<223>

<400> 11

atg gca gct gcc tgg cag gga tgg ctg ctc tgg gcc ctg ctc ctg aat	48
Met Ala Ala Ala Trp Gln Gly Trp Leu Leu Trp Ala Leu Leu Leu Asn	
1 5 10 15	
tcg gcc cag ggt gag ctc tac aca ccc act cac aaa gct ggc ttc tgc	96
Ser Ala Gln Gly Glu Leu Tyr Thr Pro Thr His Lys Ala Gly Phe Cys	
20 25 30	
acc ttt tat gaa gag tgt ggg aag aac cca gag ctt tct gga ggc ctc	144
Thr Phe Tyr Glu Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Gly Leu	
35 40 45	
aca tca cta tcc aat atc tcc tgc ttg tct aat acc cca gcc cgc cat	192
Thr Ser Leu Ser Asn Ile Ser Cys Leu Ser Asn Thr Pro Ala Arg His	
50 55 60	
gtc aca ggt gac cac ctg gct ctt ctc cag cgc gtc tgt ccc cgc cta	240
Val Thr Gly Asp His Leu Ala Leu Leu Gln Arg Val Cys Pro Arg Leu	
65 70 75 80	
tac aat ggc ccc aat gac acc tat gcc tgt tgc tct acc aag cag ctg	288
Tyr Asn Gly Pro Asn Asp Thr Tyr Ala Cys Cys Ser Thr Lys Gln Leu	
85 90 95	
gtg tca tta gac agt agc ctg tct atc acc aag gcc ctc ctt aca cgc	336
Val Ser Leu Asp Ser Ser Leu Ser Ile Thr Lys Ala Leu Leu Thr Arg	
100 105 110	
tgc ccg gca tgc tct gaa aat ttt gtg agc ata cac tgt cat aat acc	384
Cys Pro Ala Cys Ser Glu Asn Phe Val Ser Ile His Cys His Asn Thr	
115 120 125	
tgc agc cct gac cag agc ctc ttc atc aat gtt act cgc gtg gtt cag	432
Cys Ser Pro Asp Gln Ser Leu Phe Ile Asn Val Thr Arg Val Val Gln	
130 135 140	
cgg gac cct gga cag ctt cct gct gtg gtg gcc tat gag gcc ttt tat	480
Arg Asp Pro Gly Gln Leu Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr	
145 150 155 160	
caa cgc agt ttt gca gag aag gcc tat gag tcc tgt agc cgg gtg cgc	528
Gln Arg Ser Phe Ala Glu Lys Ala Tyr Glu Ser Cys Ser Arg Val Arg	
165 170 175	

atc cct gca gct gcc tcg ctg gct gtg ggc agc atg tgt gga gtg tat	576
Ile Pro Ala Ala Ala Ser Leu Ala Val Gly Ser Met Cys Gly Val Tyr	
180 185 190	
ggc tct gcc ctc tgc aat gct cag cgc tgg ctc aac ttc caa gga gac	624
Gly Ser Ala Leu Cys Asn Ala Gln Arg Trp Leu Asn Phe Gln Gly Asp	
195 200 205	
aca ggg aat ggc ctg gct ccg ctg gac atc acc ttc cac ctc ttg gag	672
Thr Gly Asn Gly Leu Ala Pro Leu Asp Ile Thr Phe His Leu Leu Glu	
210 215 220	
cct ggc cag gcc ctg gca gat ggg atg aag cca ctg gat ggg aag atc	720
Pro Gly Gln Ala Leu Ala Asp Gly Met Lys Pro Leu Asp Gly Lys Ile	
225 230 235 240	
aca ccc tgc aat gag tcc cag ggt gaa gac tcg gca gcc tgt tcc tgc	768
Thr Pro Cys Asn Glu Ser Gln Gly Glu Asp Ser Ala Ala Cys Ser Cys	
245 250 255	
cag gac tgt gca gca tcc tgc cct gtc atc cct ccg ccc ccg gcc ctg	816
Gln Asp Cys Ala Ala Ser Cys Pro Val Ile Pro Pro Pro Ala Leu	
260 265 270	
cgc cct tct ttc tac atg ggt cga atg cca ggc tgg ctg gct ctc atc	864
Arg Pro Ser Phe Tyr Met Gly Arg Met Pro Gly Trp Leu Ala Leu Ile	
275 280 285	
atc atc ttc act gct gtc ttt gta ttg ctc tct gtt gtc ctt gtg tat	912
Ile Ile Phe Thr Ala Val Phe Val Leu Leu Ser Val Val Leu Val Tyr	
290 295 300	
ctc cga gtg gct tcc aac agg aac aag aac aag aca gca ggc tcc cag	960
Leu Arg Val Ala Ser Asn Arg Asn Lys Asn Lys Thr Ala Gly Ser Gln	
305 310 315 320	
gaa gcc ccc aac ctc cct cgt aag cgc aga ttc tca cct cac act gtc	1008
Glu Ala Pro Asn Leu Pro Arg Lys Arg Arg Phe Ser Pro His Thr Val	
325 330 335	
ctt ggc cgg ttc ttc gag agc tgg gga aca agg gtg gcc tca tgg cca	1056
Leu Gly Arg Phe Phe Glu Ser Trp Gly Thr Arg Val Ala Ser Trp Pro	
340 345 350	
ctc act gtc ttg gca ctg tcc ttc ata gtt gtg ata gcc ttg tca gta	1104
Leu Thr Val Leu Ala Leu Ser Phe Ile Val Val Ile Ala Leu Ser Val	
355 360 365	
ggc ctg acc ttt ata gaa ctc acc aca gac cct gtg gaa ctg tgg tcg	1152
Gly Leu Thr Phe Ile Glu Leu Thr Thr Asp Pro Val Glu Leu Trp Ser	
370 375 380	
gcc cct aaa agc caa gcc cgg aaa gaa aag gct ttc cat gac gag cat	1200
Ala Pro Lys Ser Gln Ala Arg Lys Glu Lys Ala Phe His Asp Glu His	
385 390 395 400	
ttt ggc ccc ttc ttc cga acc aac cag att ttt gtg aca gct aag aac	1248
Phe Gly Pro Phe Phe Arg Thr Asn Gln Ile Phe Val Thr Ala Lys Asn	
405 410 415	

agg tcc agc tac aag tac gac tcc ctg ctg cta ggg ccc aag aac ttc Arg Ser Ser Tyr Lys Tyr Asp Ser Leu Leu Leu Gly Pro Lys Asn Phe 420 425 430	1296
agt ggg atc cta tcc ctg gac ttg ctg cag gag ctg ttg gag cta cag Ser Gly Ile Leu Ser Leu Asp Leu Leu Gln Glu Leu Leu Glu Leu Gln 435 440 445	1344
gag aga ctt cga cac ctg caa gtg tgg tcc cat gag gca cag cgc aac Glu Arg Leu Arg His Leu Gln Val Trp Ser His Glu Ala Gln Arg Asn 450 455 460	1392
atc tcc ctc cag gac atc tgc tat gct ccc ctc aac ccg cat aac acc Ile Ser Leu Gln Asp Ile Cys Tyr Ala Pro Leu Asn Pro His Asn Thr 465 470 475 480	1440
agc ctc act gac tgc tgt gtc aac agc ctc ctt caa tac ttc cag aac Ser Leu Thr Asp Cys Cys Val Asn Ser Leu Leu Gln Tyr Phe Gln Asn 485 490 495	1488
aac cac aca ctc ctg ctg ctc aca gcc aat cag act ctg aat ggc cag Asn His Thr Leu Leu Leu Thr Ala Asn Gln Thr Leu Asn Gly Gln 500 505 510	1536
acc tcc ctg gtg gac tgg aag gac cat ttc ctc tac tgt gcc aat gcc Thr Ser Leu Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala 515 520 525	1584
cct ctc acg tac aaa gat ggc aca gcc ctg gcc ctg agc tgc ata gct Pro Leu Thr Tyr Lys Asp Gly Thr Ala Leu Ala Leu Ser Cys Ile Ala 530 535 540	1632
gac tac ggg gca cct gtc ttc ccc ttc ctt gct gtt ggg ggc tac caa Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Val Gly Gly Tyr Gln 545 550 555 560	1680
ggg acg gac tac tcg gag gca gaa gcc ctg atc ata acc ttc tct atc Gly Thr Asp Tyr Ser Glu Ala Glu Ala Leu Ile Ile Thr Phe Ser Ile 565 570 575	1728
aat aac tac ccc gct gat gat ccc cgc atg gcc cac gcc aag ctc tgg Asn Asn Tyr Pro Ala Asp Asp Pro Arg Met Ala His Ala Lys Leu Trp 580 585 590	1776
gag gag gct ttc ttg aag gaa atg caa tcc ttc cag aga agc aca gct Glu Glu Ala Phe Leu Lys Glu Met Gln Ser Phe Gln Arg Ser Thr Ala 595 600 605	1824
gac aag ttc cag att gcg ttc tca gct gag cgt tct ctg gag gac gag Asp Lys Phe Gln Ile Ala Phe Ser Ala Glu Arg Ser Leu Glu Asp Glu 610 615 620	1872
atc aat cgc act acc atc cag gac ctg cct gtc ttt gcc atc agc tac Ile Asn Arg Thr Thr Ile Gln Asp Leu Pro Val Phe Ala Ile Ser Tyr 625 630 635 640	1920
ctt atc gtc ttc ctg tac atc tcc ctg gcc ctg ggc agc tac tcc aga Leu Ile Val Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Arg	1968

645					650					655						
tgg	agc	cga	gtt	gcg	gtg	gat	tcc	aag	gct	act	ctg	ggc	cta	ggt	ggg	2016
Trp	Ser	Arg	Val	Ala	Val	Asp	Ser	Lys	Ala	Thr	Leu	Gly	Leu	Gly	Gly	
			660					665					670			
gtg	gct	gtt	gtg	ctg	gga	gca	gtc	gtc	gct	gcc	atg	ggc	ttc	tac	tcc	2064
Val	Ala	Val	Val	Leu	Gly	Ala	Val	Val	Ala	Ala	Met	Gly	Phe	Tyr	Ser	
			675				680					685				
tac	ctg	ggt	gtc	ccc	tcc	tct	ctg	gtc	atc	att	caa	gtg	gta	cct	ttc	2112
Tyr	Leu	Gly	Val	Pro	Ser	Ser	Leu	Val	Ile	Ile	Gln	Val	Val	Pro	Phe	
	690					695					700					
ctg	gtg	ctg	gct	gtg	gga	gct	gac	aac	atc	ttc	atc	ttt	gtt	ctt	gag	2160
Leu	Val	Leu	Ala	Val	Gly	Ala	Asp	Asn	Ile	Phe	Ile	Phe	Val	Leu	Glu	
					710					715					720	
tac	cag	agg	ctg	cct	agg	atg	ccc	ggg	gag	cag	cga	gag	gct	cac	att	2208
Tyr	Gln	Arg	Leu	Pro	Arg	Met	Pro	Gly	Glu	Gln	Arg	Glu	Ala	His	Ile	
				725				730						735		
ggc	cgc	acc	ctg	ggt	agt	gtg	gcc	ccc	agc	atg	ctg	ctg	tgc	agc	ctc	2256
Gly	Arg	Thr	Leu	Gly	Ser	Val	Ala	Pro	Ser	Met	Leu	Leu	Cys	Ser	Leu	
			740					745					750			
tct	gag	gcc	atc	tgc	ttc	ttt	cta	ggg	gcc	ctg	acc	tcc	atg	cca	gct	2304
Ser	Glu	Ala	Ile	Cys	Phe	Phe	Leu	Gly	Ala	Leu	Thr	Ser	Met	Pro	Ala	
			755				760					765				
gtg	agg	acc	ttt	gcc	ttg	acc	tct	ggc	tta	gca	atc	atc	ttt	gac	ttc	2352
Val	Arg	Thr	Phe	Ala	Leu	Thr	Ser	Gly	Leu	Ala	Ile	Ile	Phe	Asp	Phe	
			770			775					780					
ctg	ctc	cag	atg	aca	gcc	ttt	gtg	gcc	ctg	ctc	tcc	ctg	gat	agc	aag	2400
Leu	Leu	Gln	Met	Thr	Ala	Phe	Val	Ala	Leu	Leu	Ser	Leu	Asp	Ser	Lys	
					790				795						800	
agg	cag	gag	gcc	tct	cgc	ccc	gac	gtc	gtg	tgc	tgc	ttt	tca	agc	cga	2448
Arg	Gln	Glu	Ala	Ser	Arg	Pro	Asp	Val	Val	Cys	Cys	Phe	Ser	Ser	Arg	
				805				810					815			
aat	ctg	ccc	cca	ccg	aaa	caa	aaa	gaa	ggc	ctc	tta	ctt	tgc	ttc	ttc	2496
Asn	Leu	Pro	Pro	Pro	Lys	Gln	Lys	Glu	Gly	Leu	Leu	Leu	Cys	Phe	Phe	
			820					825					830			
cgc	aag	ata	tac	act	ccc	ttc	ctg	ctg	cac	aga	ttc	atc	cgc	cct	gtt	2544
Arg	Lys	Ile	Tyr	Thr	Pro	Phe	Leu	Leu	His	Arg	Phe	Ile	Arg	Pro	Val	
			835				840					845				
gtg	ctg	ctg	ctc	ttt	ctg	gtc	ctg	ttt	gga	gca	aac	ctc	tac	tta	atg	2592
Val	Leu	Leu	Leu	Phe	Leu	Val	Leu	Phe	Gly	Ala	Asn	Leu	Tyr	Leu	Met	
			850			855					860					
tgc	aac	atc	agc	gtg	ggg	ctg	gac	cag	gat	ctg	gct	ctg	ccc	aag	gat	2640
Cys	Asn	Ile	Ser	Val	Gly	Leu	Asp	Gln	Asp	Leu	Ala	Leu	Pro	Lys	Asp	
					870				875						880	
tcc	tac	ctg	ata	gac	tac	ttc	ctc	ttt	ctg	aac	cgg	tac	ttg	gaa	gtg	2688

Ser Tyr Leu Ile Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Leu Glu Val	
885 890 895	
ggg cct cca gtg tac ttt gac acc acc tca ggc tac aac ttt tcc acc	2736
Gly Pro Pro Val Tyr Phe Asp Thr Thr Ser Gly Tyr Asn Phe Ser Thr	
900 905 910	
gag gca ggc atg aac gcc att tgc tct agt gca ggc tgt gag agc ttc	2784
Glu Ala Gly Met Asn Ala Ile Cys Ser Ser Ala Gly Cys Glu Ser Phe	
915 920 925	
tcc cta acc cag aaa atc cag tat gcc agt gaa ttc cct aat cag tct	2832
Ser Leu Thr Gln Lys Ile Gln Tyr Ala Ser Glu Phe Pro Asn Gln Ser	
930 935 940	
tat gtg gct att gct gca tcc tcc tgg gta gat gac ttc atc gac tgg	2880
Tyr Val Ala Ile Ala Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp	
945 950 955 960	
ctg acc cca tcc tcc tcc tgc tgc cgc att tat acc cgt ggc ccc cat	2928
Leu Thr Pro Ser Ser Ser Cys Cys Arg Ile Tyr Thr Arg Gly Pro His	
965 970 975	
aaa gat gag ttc tgt ccc tca acg gat act tcc ttc aac tgt ctc aaa	2976
Lys Asp Glu Phe Cys Pro Ser Thr Asp Thr Ser Phe Asn Cys Leu Lys	
980 985 990	
aac tgc atg aac cgc act ctg ggt ccc gtg aga ccc aca aca gaa cag	3024
Asn Cys Met Asn Arg Thr Leu Gly Pro Val Arg Pro Thr Thr Glu Gln	
995 1000 1005	
ttt cat aag tac ctg ccc tgg ttc ctg aat gat acg ccc aac atc	3069
Phe His Lys Tyr Leu Pro Trp Phe Leu Asn Asp Thr Pro Asn Ile	
1010 1015 1020	
aga tgt cct aaa ggg ggc cta gca gcg tat aga acc tct gtg aat	3114
Arg Cys Pro Lys Gly Gly Leu Ala Ala Tyr Arg Thr Ser Val Asn	
1025 1030 1035	
ttg agc tca gat ggc cag att ata gcc tcc cag ttc atg gcc tac	3159
Leu Ser Ser Asp Gly Gln Ile Ile Ala Ser Gln Phe Met Ala Tyr	
1040 1045 1050	
cac aag ccc tta cgg aac tca cag gac ttt aca gaa gct ctc cgg	3204
His Lys Pro Leu Arg Asn Ser Gln Asp Phe Thr Glu Ala Leu Arg	
1055 1060 1065	
gca tcc cgg ttg cta gca gcc aac atc aca gct gaa cta cgg aag	3249
Ala Ser Arg Leu Leu Ala Ala Asn Ile Thr Ala Glu Leu Arg Lys	
1070 1075 1080	
gtg cct ggg aca gat ccc aac ttt gag gtc ttc cct tac acg atc	3294
Val Pro Gly Thr Asp Pro Asn Phe Glu Val Phe Pro Tyr Thr Ile	
1085 1090 1095	
tcc aat gtg ttc tac cag caa tac ctg acg gtt ctc cct gag gga	3339
Ser Asn Val Phe Tyr Gln Gln Tyr Leu Thr Val Leu Pro Glu Gly	
1100 1105 1110	

atc ttc act ctt gct ctc tgc ttc gtg ccc acc ttt gtg gtc tgc Ile Phe Thr Leu Ala Leu Cys Phe Val Pro Thr Phe Val Val Cys 1115 1120 1125	3384
tac ctc cta ctg ggc ctg gac ata cgc tca ggc atc ctc aac ctg Tyr Leu Leu Leu Gly Leu Asp Ile Arg Ser Gly Ile Leu Asn Leu 1130 1135 1140	3429
ctc tcc atc att atg atc ctc gtg gac acc atc ggc ctc atg gct Leu Ser Ile Ile Met Ile Leu Val Asp Thr Ile Gly Leu Met Ala 1145 1150 1155	3474
gtg tgg ggt atc agc tac aat gct gtg tcc ctc atc aac ctt gtc Val Trp Gly Ile Ser Tyr Asn Ala Val Ser Leu Ile Asn Leu Val 1160 1165 1170	3519
acg gca gtg ggc atg tct gtg gag ttc gtg tcc cac att acc cgg Thr Ala Val Gly Met Ser Val Glu Phe Val Ser His Ile Thr Arg 1175 1180 1185	3564
tcc ttt gct gta agc acc aag cct acc cgg ctg gag aga gcc aaa Ser Phe Ala Val Ser Thr Lys Pro Thr Arg Leu Glu Arg Ala Lys 1190 1195 1200	3609
gat gct act atc ttc atg ggc agt gcg gtg ttt gct gga gtg gcc Asp Ala Thr Ile Phe Met Gly Ser Ala Val Phe Ala Gly Val Ala 1205 1210 1215	3654
atg acc aac ttc ccg ggc atc ctc atc ctg ggc ttt gct cag gcc Met Thr Asn Phe Pro Gly Ile Leu Ile Leu Gly Phe Ala Gln Ala 1220 1225 1230	3699
cag ctt atc cag att ttc ttc ttc cgc ctc aac ctc ctg atc acc Gln Leu Ile Gln Ile Phe Phe Phe Arg Leu Asn Leu Leu Ile Thr 1235 1240 1245	3744
ttg ctg ggt ctg cta cac ggc ctg gtc ttc ctg ccc gtt gtc ctc Leu Leu Gly Leu Leu His Gly Leu Val Phe Leu Pro Val Val Leu 1250 1255 1260	3789
agc tat ctg ggg cca gat gtt aac caa gct ctg gta ctg gag gag Ser Tyr Leu Gly Pro Asp Val Asn Gln Ala Leu Val Leu Glu Glu 1265 1270 1275	3834
aaa cta gcc act gag gca gcc atg gtc tca gag cct tct tgc cca Lys Leu Ala Thr Glu Ala Ala Met Val Ser Glu Pro Ser Cys Pro 1280 1285 1290	3879
cag tac ccc ttc ccg gct gat gca aac acc agt gac tat gtt aac Gln Tyr Pro Phe Pro Ala Asp Ala Asn Thr Ser Asp Tyr Val Asn 1295 1300 1305	3924
tac ggc ttt aat cca gaa ttt atc cct gaa att aat gct gct agc Tyr Gly Phe Asn Pro Glu Phe Ile Pro Glu Ile Asn Ala Ala Ser 1310 1315 1320	3969
agc tct ctg ccc aaa agt gac caa aag ttc taa Ser Ser Leu Pro Lys Ser Asp Gln Lys Phe 1325 1330	4002

<210> 12

<211> 1333

<212> PRT

<213> Mus sp.

<400> 12

Met Ala Ala Ala Trp Gln Gly Trp Leu Leu Trp Ala Leu Leu Leu Asn
1 5 10 15

Ser Ala Gln Gly Glu Leu Tyr Thr Pro Thr His Lys Ala Gly Phe Cys
20 25 30

Thr Phe Tyr Glu Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Gly Leu
35 40 45

Thr Ser Leu Ser Asn Ile Ser Cys Leu Ser Asn Thr Pro Ala Arg His
50 55 60

Val Thr Gly Asp His Leu Ala Leu Leu Gln Arg Val Cys Pro Arg Leu
65 70 75 80

Tyr Asn Gly Pro Asn Asp Thr Tyr Ala Cys Cys Ser Thr Lys Gln Leu
85 90 95

Val Ser Leu Asp Ser Ser Leu Ser Ile Thr Lys Ala Leu Leu Thr Arg
100 105 110

Cys Pro Ala Cys Ser Glu Asn Phe Val Ser Ile His Cys His Asn Thr
115 120 125

Cys Ser Pro Asp Gln Ser Leu Phe Ile Asn Val Thr Arg Val Val Gln
130 135 140

Arg Asp Pro Gly Gln Leu Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr
145 150 155 160

Gln Arg Ser Phe Ala Glu Lys Ala Tyr Glu Ser Cys Ser Arg Val Arg
165 170 175

Ile Pro Ala Ala Ala Ser Leu Ala Val Gly Ser Met Cys Gly Val Tyr

Arg Ser Ser Tyr Lys Tyr Asp Ser Leu Leu Leu Gly Pro Lys Asn Phe
420 425 430

Ser Gly Ile Leu Ser Leu Asp Leu Leu Gln Glu Leu Leu Glu Leu Gln
435 440 445

Glu Arg Leu Arg His Leu Gln Val Trp Ser His Glu Ala Gln Arg Asn
450 455 460

Ile Ser Leu Gln Asp Ile Cys Tyr Ala Pro Leu Asn Pro His Asn Thr
465 470 475 480

Ser Leu Thr Asp Cys Cys Val Asn Ser Leu Leu Gln Tyr Phe Gln Asn
485 490 495

Asn His Thr Leu Leu Leu Leu Thr Ala Asn Gln Thr Leu Asn Gly Gln
500 505 510

Thr Ser Leu Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala
515 520 525

Pro Leu Thr Tyr Lys Asp Gly Thr Ala Leu Ala Leu Ser Cys Ile Ala
530 535 540

Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Val Gly Gly Tyr Gln
545 550 555 560

Gly Thr Asp Tyr Ser Glu Ala Glu Ala Leu Ile Ile Thr Phe Ser Ile
565 570 575

Asn Asn Tyr Pro Ala Asp Asp Pro Arg Met Ala His Ala Lys Leu Trp
580 585 590

Glu Glu Ala Phe Leu Lys Glu Met Gln Ser Phe Gln Arg Ser Thr Ala
595 600 605

Asp Lys Phe Gln Ile Ala Phe Ser Ala Glu Arg Ser Leu Glu Asp Glu
610 615 620

Ile Asn Arg Thr Thr Ile Gln Asp Leu Pro Val Phe Ala Ile Ser Tyr
625 630 635 640

Leu Ile Val Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Arg
645 650 655

Trp Ser Arg Val Ala Val Asp Ser Lys Ala Thr Leu Gly Leu Gly Gly
660 665 670

Val Ala Val Val Leu Gly Ala Val Val Ala Ala Met Gly Phe Tyr Ser
675 680 685

Tyr Leu Gly Val Pro Ser Ser Leu Val Ile Ile Gln Val Val Pro Phe
690 695 700

Leu Val Leu Ala Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu
705 710 715 720

Tyr Gln Arg Leu Pro Arg Met Pro Gly Glu Gln Arg Glu Ala His Ile
725 730 735

Gly Arg Thr Leu Gly Ser Val Ala Pro Ser Met Leu Leu Cys Ser Leu
740 745 750

Ser Glu Ala Ile Cys Phe Phe Leu Gly Ala Leu Thr Ser Met Pro Ala
755 760 765

Val Arg Thr Phe Ala Leu Thr Ser Gly Leu Ala Ile Ile Phe Asp Phe
770 775 780

Leu Leu Gln Met Thr Ala Phe Val Ala Leu Leu Ser Leu Asp Ser Lys
785 790 795 800

Arg Gln Glu Ala Ser Arg Pro Asp Val Val Cys Cys Phe Ser Ser Arg
805 810 815

Asn Leu Pro Pro Pro Lys Gln Lys Glu Gly Leu Leu Leu Cys Phe Phe
820 825 830

Arg Lys Ile Tyr Thr Pro Phe Leu Leu His Arg Phe Ile Arg Pro Val
835 840 845

Val Leu Leu Leu Phe Leu Val Leu Phe Gly Ala Asn Leu Tyr Leu Met
850 855 860

Cys Asn Ile Ser Val Gly Leu Asp Gln Asp Leu Ala Leu Pro Lys Asp
865 870 875 880

Ser Tyr Leu Ile Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Leu Glu Val
885 890 895

Gly Pro Pro Val Tyr Phe Asp Thr Thr Ser Gly Tyr Asn Phe Ser Thr
900 905 910

Glu Ala Gly Met Asn Ala Ile Cys Ser Ser Ala Gly Cys Glu Ser Phe
915 920 925

Ser Leu Thr Gln Lys Ile Gln Tyr Ala Ser Glu Phe Pro Asn Gln Ser
930 935 940

Tyr Val Ala Ile Ala Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp
945 950 955 960

Leu Thr Pro Ser Ser Ser Cys Cys Arg Ile Tyr Thr Arg Gly Pro His
965 970 975

Lys Asp Glu Phe Cys Pro Ser Thr Asp Thr Ser Phe Asn Cys Leu Lys
980 985 990

Asn Cys Met Asn Arg Thr Leu Gly Pro Val Arg Pro Thr Thr Glu Gln
995 1000 1005

Phe His Lys Tyr Leu Pro Trp Phe Leu Asn Asp Thr Pro Asn Ile
1010 1015 1020

Arg Cys Pro Lys Gly Gly Leu Ala Ala Tyr Arg Thr Ser Val Asn
1025 1030 1035

Leu Ser Ser Asp Gly Gln Ile Ile Ala Ser Gln Phe Met Ala Tyr
1040 1045 1050

His Lys Pro Leu Arg Asn Ser Gln Asp Phe Thr Glu Ala Leu Arg
1055 1060 1065

Ala Ser Arg Leu Leu Ala Ala Asn Ile Thr Ala Glu Leu Arg Lys
1070 1075 1080

Val Pro Gly Thr Asp Pro Asn Phe Glu Val Phe Pro Tyr Thr Ile
1085 1090 1095

Ser Asn Val Phe Tyr Gln Gln Tyr Leu Thr Val Leu Pro Glu Gly
1100 1105 1110

Ile Phe Thr Leu Ala Leu Cys Phe Val Pro Thr Phe Val Val Cys

1115	1120	1125
Tyr Leu Leu Leu Gly Leu Asp 1130	Ile Arg Ser Gly Ile 1135	Leu Asn Leu 1140
Leu Ser Ile Ile Met Ile Leu 1145	Val Asp Thr Ile 1150	Gly Leu Met Ala 1155
Val Trp Gly Ile Ser Tyr Asn 1160	Ala Val Ser Leu 1165	Ile Asn Leu Val 1170
Thr Ala Val Gly Met Ser Val 1175	Glu Phe Val Ser 1180	His Ile Thr Arg 1185
Ser Phe Ala Val Ser Thr Lys 1190	Pro Thr Arg Leu 1195	Glu Arg Ala Lys 1200
Asp Ala Thr Ile Phe Met Gly 1205	Ser Ala Val Phe 1210	Ala Gly Val Ala 1215
Met Thr Asn Phe Pro Gly Ile 1220	Leu Ile Leu Gly Phe 1225	Ala Gln Ala 1230
Gln Leu Ile Gln Ile Phe Phe 1235	Phe Arg Leu Asn 1240	Leu Leu Ile Thr 1245
Leu Leu Gly Leu Leu His Gly 1250	Leu Val Phe Leu 1255	Pro Val Val Leu 1260
Ser Tyr Leu Gly Pro Asp Val 1265	Asn Gln Ala Leu 1270	Val Leu Glu Glu 1275
Lys Leu Ala Thr Glu Ala Ala 1280	Met Val Ser Glu 1285	Pro Ser Cys Pro 1290
Gln Tyr Pro Phe Pro Ala Asp 1295	Ala Asn Thr Ser 1300	Asp Tyr Val Asn 1305
Tyr Gly Phe Asn Pro Glu Phe 1310	Ile Pro Glu Ile 1315	Asn Ala Ala Ser 1320
Ser Ser Leu Pro Lys Ser Asp 1325	Gln Lys Phe 1330	

<210> 13

<211> 3999

<212> DNA

<213> Mus sp.

<220>

<221> misc_feature

<222> (1)..(3999)

<223> n is g or a or t or c

<400> 13

atggcngcng cntggcargg ntggytnytn tgggcnytny tnytnaayws ngcncarggn	60
garytnataya cncnacnca yaargcnggn ttytgyacnt tytaygarga rtgyggnaar	120
aayccngary tnwsnggngg nytnacnwsn ytnwsnaaya thwsntgyyt nwsnaayacn	180
ccngcnmgnc aygtnacngg ngaycayytn gcnytnytnc armngtntg yccnmgnytn	240
tayaayggnc cnaaygayac ntaygcntgy tgywsnacna arcarytngt nwsnytngay	300
wsnwsnytnw snathacnaa rgcnynytn acnmngtgyc cngcntgyws ngaraaytty	360
gtwnsnathc aytgycayaa yacntgywsn ccngaycarw snytnnttyat haaygtnacn	420
mgngtngtnc armnggaycc nggncarytn ccngcngtng tngcntayga rgcnttytay	480
carmgnwsnt tygcngaraa rgcntaygar wsntgywsnm gngtnmgntat hccngcngcn	540
gcnwsnytnng cngtnggnws natgtgyggg gtntayggngw sngcnytnng yaaygcncar	600
mgntggytna ayttycargg ngayacnggn aayggnytnng cncnytnga yathacntty	660
cayytnytnng arccnggnca rgcnynngcn gayggntatga arccnytnga yggnaarath	720
acncntgya aygarwsnca rggngargay wsngcngcnt gywsntgyca rgaytgygcn	780
gcnwsntgyc cngtnathcc nccncncncn gcnytnmgnc cnwsnttyta yatgggnmg	840
atgccngnt ggytnngcnyt nathathath ttyacngcng tnttygtnyt nytnwsngtn	900
gtnytngtnt ayytnmgnt ngcnwsnaay mgnaayaara ayaaracngc nggnwsncar	960
gargcnccna ayytnccnmg naarmgnmgnt ttywsnccnc ayacngtnyt nggnmgntty	1020
ttygarwsnt ggggnacnmg ngtnngcnwsn tggcnytna cngtnytnge nytnwsntty	1080
athgtngtna thgcnytnws ngtnngnytn acnttyathg arytnacnac ngayccngtn	1140
garytnntggw sngcnccnaa rwsncargcn mgnaargara argcnttyca ygaygar'cay	1200
ttyggncnt tyttymgnac naaycarath ttygtnacng cnaaraaymg nwsnwsntay	1260

aartaygayw snytnytnyt nggnccnaar aayttywsng gnathytnws nytngayytn 1320
ytncargary tnytngaryt ncargarmgn ytnmgncayy tncargtntg gwsncaygar 1380
gcncarmgna ayathwsnyt ncargayath tgytaygcnc cnytnaaycc ncayaayacn 1440
wsnytnacng aytgytgygt naaywsnytn ytncartayt tycaraayaa ycayacnytn 1500
ytntytnytna cngcnaayca racnytnaay ggncaracnw snytngtnga ytggaaargay 1560
cayttyytnt aytgygcnaa ygcncnytn acntayaarg ayggnacngc nytngcnytn 1620
wsntgyathg cngaytaygg ngcncngtn ttyccnttyy tngcngtngg nggntaycar 1680
ggnacngayt aywsngargc ngargcnytn athathacnt tywsnathaa yaaytayccn 1740
gengaygayc cnmgntatggc ncaygcnaar ytntgggarg argcnttyyt naargaratg 1800
carwsnttyc armgnwsnac ngcngayaar ttycarathg cnttywsngc ngarmgnwsn 1860
ytngargayg arathaaymg nacnacnath cargayytn cngtnttygc nathwsntay 1920
ytnathgtnt tyyntayay hwsnytnngcn ytnggnwsnt aywsnmngntg gwsnmngntn 1980
gngtngayw snaargcnac nytnggnytn ggnggngtng cngtngtnyt nggngcngtn 2040
gtngcngcna tgggnttyta ywsntayytn ggngtncnw snwsnytngt nathathcar 2100
gtngtncnt tyytngtnyt ngcngtnggn gngayaaya thttyathht ygtnytngar 2160
taycarmgny tnccnmgnat gccngngar carmgngarg cncayathgg nmgnacnytn 2220
ggwngtng cncnwsnat gytnytnygy wsnytnwsng argcnathgt yttytyytn 2280
ggngcnytna cnwsnatgcc ngcngtnmgn acnttygcny tnacnwsngg nytngcnath 2340
athttygayt tyytnytnca ratgaengcn ttygtngcny tnytnwsnyt ngaywsnaar 2400
mgncargarg cnwsnmgncc ngaygtngtn tgytgytyw snwsnmgnaa yytnccnccn 2460
ccnaarcara argarggnyt nytnytnygy ttyttymgna arathtayac nccnttyytn 2520
ytncaymgnt tyathmgnc ngtngtnytn ytntynttyy tngtnytnntt yggngcnaay 2580
ytntayytna tgtgyaayat hwsngtnggn ytngaycarg ayytngcnytn nccnaargay 2640
wsntayytna thgaytaytt yytnnttyytn aaymgntayy tngargtngg nccncngtn 2700
tayttygaya cnacnwsngg ntayaaytty wsnacngarg cnggnatgaa ygcnathtgy 2760
wsnwsngcng gntgygarws nttywsnytn acncaraara thcartaygc nwsngartty 2820
ccnaaycarw sntaygtngc nathgcngcn wsnwsntggg tngaygaytt yathgaytgg 2880
ytnacnccnw snwsnwsntg ytgymgnath tayacnmng gncncayaa rgaygartty 2940
tgyccnwsna cngayacnws nttyaaytgy ytnaaraayt gyatgaaymg nacnytnngn 3000

ccngtnmgnc cna cna cnga rcarttycay aartayytnc cntgggttyt naaygayacn 3060
ccnaayathm gnt gyccnaa rggnggnytn gcngcntaym gnacnwsngt naayytwnsn 3120
wsngayggnc arathathgc nwsncartty atggcntayc ayaarccnyt nmgnaywsn 3180
cargayttya cngargcnyt nmngcnwsn mgnytnytng cngcnaayat hacngcngar 3240
ytnmgnaarg tnc cnggnac ngayccnaay ttygargtnt tyccntayac nathwsnaay 3300
gtnttytayc arcartayyt nacngtnytn ccngarggna thttyacnyt ngcnytnygy 3360
ttygtncna cnt tygtngt ntgytayyt ytnynggny tngayathmg nwsnggnath 3420
ytnaayytny tnwsnathat hatgathytn gtngayacna thggnytnat ggngtntgg 3480
ggngnathwsnt ayaaygcngt nwsnytnath aayytngtna cngcngtngg natgwsngtn 3540
garttygtnw snayathac nmgnwsntty gcngtnwsna cnaarccnac nmgnytngar 3600
mgngcnaarg aygcnaacnat httyatgggn wsngcngtnt tygcnggngt ngcnatgacn 3660
aayttyccng gnathytnat hytnggntty gcncargcnc arytnathca rathttytty 3720
tymgnytna ayytnytnat hacnytnytn ggnytnytn cnyggnytngt nttytnccn 3780
gtngtnytnw sntayytngg nccngaygt n aaycargcny tngtnytna rgaraarytn 3840
gcnaacngarg cngcnatggt nwsngarccn wsntgyccnc artayccntt yccngcngay 3900
gcnaayacnw sngaytaygt naaytaygg n ttyaayccng arttyathcc ngarathaay 3960
gcngcnwsnw snwsnytncc naarwsngay caraartty 3999

<210> 14

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 14

tcttcaccct tgctctttgc

20

<210> 15

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 15

aatgatggag agtaggttga ggat

24

<210> 16

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 16

tgcccacctt tgttgtctgc taccta

26

<210> 17

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 17

atcgctgaca ggatgcagaa g

21

<210> 18

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 18

tcaggaggag caatgatctt ga

22

<210> 19

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 19

agattactgc cctggctcct agcaccatta

30

<210> 20

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 20

atcctcatcc tgggctttgc

20

<210> 21

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 21

gcaaggtgat caggaggttg a

21

<210> 22

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 22

cccagcttat ccagattttc ttcttccgc

29

<210> 23

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 23

tcttcaccct tgctctttgc

20

<210> 24

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 24

aatgatggag agtaggttga ggat

24

<210> 25

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 25
tgcccacctt tgttgtctgc tacc

24

<210> 26

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 26
agcacctgtc cactgaagat ttc

23

<210> 27

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 27
tggacgctga gcttcagttc t

21

<210> 28

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 28

cttctctgcg ctgcctcgat ggaa

24

<210> 29

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 29

agtaaaaagg gctcgcagga t

21

<210> 30

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 30

ggcagctggt gacatcagag a

21

<210> 31

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 31
aggaggccat gcaggcctac tctga

25

<210> 32

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 32
gagtccacgg tcagtccatg t

21

<210> 33

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 33
ttatgaacaa caatgccaag caa

23

<210> 34

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 34
agtccttagg tagtggctta gtccctggaa gctc

34

<210> 35

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> probe

<400> 35

gtaatacgac tcactatagg gccctgacgg tccttcctga gggaatcttc ac

52

<210> 36

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> probe

<400> 36

gtaatacgac tcactatagg gcctgggaag ttggatcatgg ccactccagc

50

<210> 37

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> FLAG tag

<400> 37

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 38

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> motif

<400> 38

Tyr Gln Arg Leu
1

<210> 39

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> antigen

<400> 39

Glu Gln Phe His Lys Tyr Leu Pro Trp Phe Leu Asn Asp Pro Pro Asn
1 5 10 15

Ile Arg Cys

<210> 40

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> antigen

<400> 40

Glu Ala Phe Tyr Gln Arg Ser Phe Ala Glu Lys Ala Tyr Glu Ser Cys
1 5 10 15

<210> 41

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> antigen

<400> 41

Gly Gln Thr Ser Leu Val Asp Trp Lys Asp His Phe Leu Tyr Cys
1 5 10 15

<210> 42

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> antigen

<400> 42

Cys Ala Asn Ala Pro Leu Thr Phe Lys Asp Gly Thr Ala Leu Ala Leu
1 5 10 15

Ser

<210> 43

<211> 5092

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (57)..(4136)

<223>

<400> 43

cttggctggtt cctgaggcct ggcctggctc cccgctgacc ccttcccaga cctggg atg	59
Met	
1	
gcg gag gcc ggc ctg agg ggc tgg ctg ctg tgg gcc ctg ctc ctg cgc	107
Ala Glu Ala Gly Leu Arg Gly Trp Leu Leu Trp Ala Leu Leu Leu Arg	
5 10 15	
ttg gcc cag agt gag cct tac aca acc atc cac cag cct ggc tac tgc	155
Leu Ala Gln Ser Glu Pro Tyr Thr Thr Ile His Gln Pro Gly Tyr Cys	
20 25 30	
gcc ttc tat gac gaa tgt ggg aag aac cca gag ctg tct gga agc ctc	203
Ala Phe Tyr Asp Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Ser Leu	
35 40 45	
atg aca ctc tcc aac gtg tcc tgc ctg tcc aac acg ccg gcc cgc aag	251
Met Thr Leu Ser Asn Val Ser Cys Leu Ser Asn Thr Pro Ala Arg Lys	
50 55 60 65	
atc aca ggt gat cac ctg atc cta tta cag aag atc tgc ccc cgc ctc	299
Ile Thr Gly Asp His Leu Ile Leu Leu Gln Lys Ile Cys Pro Arg Leu	
70 75 80	
tac acc ggc ccc aac acc caa gcc tgc tgc tcc gcc aag cag ctg gta	347
Tyr Thr Gly Pro Asn Thr Gln Ala Cys Cys Ser Ala Lys Gln Leu Val	
85 90 95	
tca ctg gaa gcg agt ctg tcg atc acc aag gcc ctc ctc acc cgc tgc	395
Ser Leu Glu Ala Ser Leu Ser Ile Thr Lys Ala Leu Leu Thr Arg Cys	
100 105 110	
cca gcc tgc tct gac aat ttt gtg aac ctg cac tgc cac aac acg tgc	443
Pro Ala Cys Ser Asp Asn Phe Val Asn Leu His Cys His Asn Thr Cys	
115 120 125	
agc ccc aat cag agc ctc ttc atc aat gtg acc cgc gtg gcc cag cta	491
Ser Pro Asn Gln Ser Leu Phe Ile Asn Val Thr Arg Val Ala Gln Leu	
130 135 140 145	
ggg gct gga caa ctc cca gct gtg gtg gcc tat gag gcc ttc tac cag	539
Gly Ala Gly Gln Leu Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr Gln	
150 155 160	
cat agc ttt gcc gag cag agc tat gac tcc tgc agc cgt gtg cgc gtc	587
His Ser Phe Ala Glu Gln Ser Tyr Asp Ser Cys Ser Arg Val Arg Val	
165 170 175	
cct gca gct gcc acg ctg gct gtg ggc acc atg tgt ggc gtg tat ggc	635
Pro Ala Ala Ala Thr Leu Ala Val Gly Thr Met Cys Gly Val Tyr Gly	
180 185 190	
tct gcc ctt tgc aat gcc cag cgc tgg ctc aac ttc cag gga gac aca	683

Ser	Ala	Leu	Cys	Asn	Ala	Gln	Arg	Trp	Leu	Asn	Phe	Gln	Gly	Asp	Thr		
195						200					205						
ggc	aat	ggt	ctg	gcc	cca	ctg	gac	atc	acc	ttc	cac	ctc	ttg	gag	cct		731
Gly	Asn	Gly	Leu	Ala	Pro	Leu	Asp	Ile	Thr	Phe	His	Leu	Leu	Glu	Pro		
210					215					220					225		
ggc	cag	gcc	gtg	ggg	agt	ggg	att	cag	cct	ctg	aat	gag	ggg	gtt	gca		779
Gly	Gln	Ala	Val	Gly	Ser	Gly	Ile	Gln	Pro	Leu	Asn	Glu	Gly	Val	Ala		
				230					235					240			
cgt	tgc	aat	gag	tcc	caa	ggt	gac	gac	gtg	gcg	acc	tgc	tcc	tgc	caa		827
Arg	Cys	Asn		Glu	Ser	Gln	Gly	Asp	Asp	Val	Ala	Thr	Cys	Ser	Cys	Gln	
				245				250						255			
gac	tgt	gct	gca	tcc	tgt	cct	gcc	ata	gcc	cgc	ccc	cag	gcc	ctc	gac		875
Asp	Cys	Ala	Ala	Ser	Cys	Pro	Ala	Ile	Ala	Arg	Pro	Gln	Ala	Leu	Asp		
		260					265					270					
tcc	acc	ttc	tac	ctg	ggc	cag	atg	ccg	ggc	agt	ctg	gtc	ctc	atc	atc		923
Ser	Thr	Phe	Tyr	Leu	Gly	Gln	Met	Pro	Gly	Ser	Leu	Val	Leu	Ile	Ile		
	275					280					285						
atc	ctc	tgc	tct	gtc	ttc	gct	gtg	gtc	acc	atc	ctg	ctt	gtg	gga	ttc		971
Ile	Leu	Cys	Ser	Val	Phe	Ala	Val	Val	Thr	Ile	Leu	Leu	Val	Gly	Phe		
290					295				300					305			
cgt	gtg	gcc	ccc	gcc	agg	gac	aaa	agc	aag	atg	gtg	gac	ccc	aag	aag		1019
Arg	Val	Ala	Pro	Ala	Arg	Asp	Lys	Ser	Lys	Met	Val	Asp	Pro	Lys	Lys		
				310				315						320			
ggc	acc	agc	ctc	tct	gac	aag	ctc	agc	ttc	tcc	acc	cac	acc	ctc	ctt		1067
Gly	Thr	Ser	Leu	Ser	Asp	Lys	Leu	Ser	Phe	Ser	Thr	His	Thr	Leu	Leu		
			325					330					335				
ggc	cag	ttc	ttc	cag	ggc	tgg	ggc	acg	tgg	gtg	gct	tcg	tgg	cct	ctg		1115
Gly	Gln	Phe	Phe	Gln	Gly	Trp	Gly	Thr	Trp	Val	Ala	Ser	Trp	Pro	Leu		
		340				345						350					
acc	atc	ttg	gtg	cta	tct	gtc	atc	ccg	gtg	gtg	gcc	ttg	gca	gcg	ggc		1163
Thr	Ile	Leu	Val	Leu	Ser	Val	Ile	Pro	Val	Val	Ala	Leu	Ala	Ala	Gly		
	355					360					365						
ctg	gtc	ttt	aca	gaa	ctc	act	acg	gac	ccc	gtg	gag	ctg	tgg	tcg	gcc		1211
Leu	Val	Phe	Thr	Glu	Leu	Thr	Thr	Asp	Pro	Val	Glu	Leu	Trp	Ser	Ala		
370					375					380					385		
ccc	aac	agc	caa	gcc	cgg	agt	gag	aaa	gct	ttc	cat	gac	cag	cat	ttc		1259
Pro	Asn	Ser	Gln	Ala	Arg	Ser	Glu	Lys	Ala	Phe	His	Asp	Gln	His	Phe		
				390				395						400			
ggc	ccc	ttc	ttc	cga	acc	aac	cag	gtg	atc	ctg	acg	gct	cct	aac	cgg		1307
Gly	Pro	Phe	Phe	Arg	Thr	Asn	Gln	Val	Ile	Leu	Thr	Ala	Pro	Asn	Arg		
			405					410					415				
tcc	agc	tac	agg	tat	gac	tct	ctg	ctg	ctg	ggg	ccc	aag	aac	ttc	agc		1355
Ser	Ser	Tyr	Arg	Tyr	Asp	Ser	Leu	Leu	Leu	Gly	Pro	Lys	Asn	Phe	Ser		
		420					425					430					

gga atc ctg gac ctg gac ttg ctg ctg gag ctg cta gag ctg cag gag Gly Ile Leu Asp Leu Asp Leu Leu Leu Glu Leu Leu Glu Leu Gln Glu 435 440 445	1403
agg ctg cgg cac ctc cag gta tgg tgc ccc gaa gca cag cgc aac atc Arg Leu Arg His Leu Gln Val Trp Ser Pro Glu Ala Gln Arg Asn Ile 450 455 460 465	1451
tcc ctg cag gac atc tgc tac gcc ccc ctc aat ccg gac aat acc agt Ser Leu Gln Asp Ile Cys Tyr Ala Pro Leu Asn Pro Asp Asn Thr Ser 470 475 480	1499
ctc tac gac tgc tgc atc aac agc ctc ctg cag tat ttc cag aac aac Leu Tyr Asp Cys Cys Ile Asn Ser Leu Leu Gln Tyr Phe Gln Asn Asn 485 490 495	1547
cgc acg ctc ctg ctg ctc aca gcc aac cag aca ctg atg ggg cag acc Arg Thr Leu Leu Leu Leu Thr Ala Asn Gln Thr Leu Met Gly Gln Thr 500 505 510	1595
tcc caa gtc gac tgg aag gac cat ttt ctg tac tgt gcc aat gcc ccg Ser Gln Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala Pro 515 520 525	1643
ctc acc ttc aag gat ggc aca gcc ctg gcc ctg agc tgc atg gct gac Leu Thr Phe Lys Asp Gly Thr Ala Leu Ala Leu Ser Cys Met Ala Asp 530 535 540 545	1691
tac ggg gcc cct gtc ttc ccc ttc ctt gcc att ggg ggg tac aaa gga Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Ile Gly Gly Tyr Lys Gly 550 555 560	1739
aag gac tat tct gag gca gag gcc ctg atc atg acg ttc tcc ctc aac Lys Asp Tyr Ser Glu Ala Glu Ala Leu Ile Met Thr Phe Ser Leu Asn 565 570 575	1787
aat tac cct gcc ggg gac ccc cgt ctg gcc cag gcc aag ctg tgg gag Asn Tyr Pro Ala Gly Asp Pro Arg Leu Ala Gln Ala Lys Leu Trp Glu 580 585 590	1835
gag gcc ttc tta gag gaa atg cga gcc ttc cag cgt cgg atg gct ggc Glu Ala Phe Leu Glu Glu Met Arg Ala Phe Gln Arg Arg Met Ala Gly 595 600 605	1883
atg ttc cag gtc acg ttc atg gct gag cgc tct ctg gaa gac gag atc Met Phe Gln Val Thr Phe Met Ala Glu Arg Ser Leu Glu Asp Glu Ile 610 615 620 625	1931
aat cgc acc aca gct gaa gac ctg ccc atc ttt gcc acc agc tac att Asn Arg Thr Thr Ala Glu Asp Leu Pro Ile Phe Ala Thr Ser Tyr Ile 630 635 640	1979
gtc ata ttc ctg tac atc tct ctg gcc ctg ggc agc tat tcc agc tgg Val Ile Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Ser Trp 645 650 655	2027
agc cga gtg atg gtg gac tcc aag gcc acg ctg ggc ctc ggc ggg gtg Ser Arg Val Met Val Asp Ser Lys Ala Thr Leu Gly Leu Gly Gly Val 660 665 670	2075

gcc gtg gtc ctg gga gca gtc atg gct gcc atg ggc ttc ttc tcc tac Ala Val Val Leu Gly Ala Val Met Ala Ala Met Gly Phe Phe Ser Tyr 675 680 685	2123
ttg ggt atc cgc tcc tcc ctg gtc atc ctg caa gtg gtt cct ttc ctg Leu Gly Ile Arg Ser Ser Leu Val Ile Leu Gln Val Val Pro Phe Leu 690 695 700 705	2171
gtg ctg tcc gtg ggg gct gat aac atc ttc atc ttt gtt ctc gag tac Val Leu Ser Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu Tyr 710 715 720	2219
cag agg ctg ccc cgg agg cct ggg gag cca cga gag gtc cac att ggg Gln Arg Leu Pro Arg Arg Pro Gly Glu Pro Arg Glu Val His Ile Gly 725 730 735	2267
cga gcc cta ggc agg gtg gct ccc agc atg ctg ttg tgc agc ctc tct Arg Ala Leu Gly Arg Val Ala Pro Ser Met Leu Leu Cys Ser Leu Ser 740 745 750	2315
gag gcc atc tgc ttc ttc cta ggg gcc ctg acc ccc atg cca gct gtg Glu Ala Ile Cys Phe Phe Leu Gly Ala Leu Thr Pro Met Pro Ala Val 755 760 765	2363
cgg acc ttt gcc ctg acc tct ggc ctt gca gtg atc ctt gac ttc ctc Arg Thr Phe Ala Leu Thr Ser Gly Leu Ala Val Ile Leu Asp Phe Leu 770 775 780 785	2411
ctg cag atg tca gcc ttt gtg gcc ctg ctc tcc ctg gac agc aag agg Leu Gln Met Ser Ala Phe Val Ala Leu Leu Ser Leu Asp Ser Lys Arg 790 795 800	2459
cag gag gcc tcc cgg ttg gac gtc tgc tgc tgt gtc aag ccc cag gag Gln Glu Ala Ser Arg Leu Asp Val Cys Cys Cys Val Lys Pro Gln Glu 805 810 815	2507
ctg ccc ccg cct ggc cag gga gag ggg ctc ctg ctt ggc ttc ttc caa Leu Pro Pro Pro Gly Gln Gly Glu Gly Leu Leu Leu Gly Phe Phe Gln 820 825 830	2555
aag gct tat gcc ccc ttc ctg ctg cac tgg atc act cga ggt gtt gtg Lys Ala Tyr Ala Pro Phe Leu Leu His Trp Ile Thr Arg Gly Val Val 835 840 845	2603
ctg ctg ctg ttt ctc gcc ctg ttc gga gtg agc ctc tac tcc atg tgc Leu Leu Leu Phe Leu Ala Leu Phe Gly Val Ser Leu Tyr Ser Met Cys 850 855 860 865	2651
cac atc agc gtg gga ctg gac cag gag ctg gcc ctg ccc aag gac tcg His Ile Ser Val Gly Leu Asp Gln Glu Leu Ala Leu Pro Lys Asp Ser 870 875 880	2699
tac ctg ctt gac tat ttc ctc ttt ctg aac cgc tac ttc gag gtg ggg Tyr Leu Leu Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Phe Glu Val Gly 885 890 895	2747
gcc ccg gtg tac ttt gtt acc acc ttg ggc tac aac ttc tcc agc gag Ala Pro Val Tyr Phe Val Thr Thr Leu Gly Tyr Asn Phe Ser Ser Glu	2795

900	905	910	
gct ggg atg aat gcc atc tgc tcc agt gca ggc tgc aac aac ttc tcc Ala Gly Met Asn Ala Ile Cys Ser Ser Ala Gly Cys Asn Asn Phe Ser 915 920 925			2843
ttc acc cag aag atc cag tat gcc aca gag ttc cct gag cag tct tac Phe Thr Gln Lys Ile Gln Tyr Ala Thr Glu Phe Pro Glu Gln Ser Tyr 930 935 940 945			2891
ctg gcc atc cct gcc tcc tcc tgg gtg gat gac ttc att gac tgg ctg Leu Ala Ile Pro Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp Leu 950 955 960			2939
acc ccg tcc tcc tgc tgc cgc ctt tat ata tct ggc ccc aat aag gac Thr Pro Ser Ser Cys Cys Arg Leu Tyr Ile Ser Gly Pro Asn Lys Asp 965 970 975			2987
aag ttc tgc ccc tcg acc gtc aac tct ctg aac tgc cta aag aac tgc Lys Phe Cys Pro Ser Thr Val Asn Ser Leu Asn Cys Leu Lys Asn Cys 980 985 990			3035
atg agc atc acg atg ggc tct gtg agg ccc tcg gtg gag cag ttc cat Met Ser Ile Thr Met Gly Ser Val Arg Pro Ser Val Glu Gln Phe His 995 1000 1005			3083
aag tat ctt ccc tgg ttc ctg aac gac cgg ccc aac atc aaa tgt Lys Tyr Leu Pro Trp Phe Leu Asn Asp Arg Pro Asn Ile Lys Cys 1010 1015 1020			3128
ccc aaa ggc ggc ctg gca gca tac agc acc tct gtg aac ttg act Pro Lys Gly Gly Leu Ala Ala Tyr Ser Thr Ser Val Asn Leu Thr 1025 1030 1035			3173
tca gat ggc cag gtt tta gac aca gtt gcc att ctg tca ccc agg Ser Asp Gly Gln Val Leu Asp Thr Val Ala Ile Leu Ser Pro Arg 1040 1045 1050			3218
ctg gag tac agt ggc aca atc tcg gct cac tgc aac ctc tac ctc Leu Glu Tyr Ser Gly Thr Ile Ser Ala His Cys Asn Leu Tyr Leu 1055 1060 1065			3263
ctg gat tca gcc tcc agg ttc atg gcc tat cac aag ccc ctg aaa Leu Asp Ser Ala Ser Arg Phe Met Ala Tyr His Lys Pro Leu Lys 1070 1075 1080			3308
aac tca cag gat tac aca gaa gct ctg cgg gca gct cga gag ctg Asn Ser Gln Asp Tyr Thr Glu Ala Leu Arg Ala Ala Arg Glu Leu 1085 1090 1095			3353
gca gcc aac atc act gct gac ctg cgg aaa gtg cct gga aca gac Ala Ala Asn Ile Thr Ala Asp Leu Arg Lys Val Pro Gly Thr Asp 1100 1105 1110			3398
ccg gct ttt gag gtc ttc ccc tac acg atc acc aat gtg ttt tat Pro Ala Phe Glu Val Phe Pro Tyr Thr Ile Thr Asn Val Phe Tyr 1115 1120 1125			3443
gag cag tac ctg acc atc ctc cct gag ggg ctc ttc atg ctc agc			3488

Glu 1130	Gln	Tyr	Leu	Thr	Ile 1135	Leu	Pro	Glu	Gly	Leu 1140	Phe	Met	Leu	Ser	
ctc 1145	tgc	ctt	gtg	ccc	acc 1150	ttc	gct	gtc	tcc	tgc 1155	ctc	ctg	ctg	ggc	3533
Leu	Cys	Leu	Val	Pro	Thr	Phe	Ala	Val	Ser	Cys	Leu	Leu	Leu	Gly	
ctg 1160	gac	ctg	cgc	tcc	ggc 1165	ctc	ctc	aac	ctg	ctc 1170	tcc	att	gtc	atg	3578
Leu	Asp	Leu	Arg	Ser	Gly	Leu	Leu	Asn	Leu	Leu	Ser	Ile	Val	Met	
atc 1175	ctc	gtg	gac	act	gtc 1180	ggc	ttc	atg	gcc	ctg 1185	tgg	ggc	atc	agt	3623
Ile	Leu	Val	Asp	Thr	Val	Gly	Phe	Met	Ala	Leu	Trp	Gly	Ile	Ser	
tac 1190	aat	gct	gtg	tcc	ctc 1195	atc	aac	ctg	gtc	tgc 1200	gcg	gtg	ggc	atg	3668
Tyr	Asn	Ala	Val	Ser	Leu	Ile	Asn	Leu	Val	Ser	Ala	Val	Gly	Met	
tct 1205	gtg	gag	ttt	gtg	tcc 1210	cac	att	acc	cgc	tcc 1215	ttt	gcc	atc	agc	3713
Ser	Val	Glu	Phe	Val	Ser	His	Ile	Thr	Arg	Ser	Phe	Ala	Ile	Ser	
acc 1220	aag	ccc	acc	tgg	ctg 1225	gag	agg	gcc	aaa	gag 1230	gcc	acc	atc	tct	3758
Thr	Lys	Pro	Thr	Trp	Leu	Glu	Arg	Ala	Lys	Glu	Ala	Thr	Ile	Ser	
atg 1235	gga	agt	gcg	gtg	ttt 1240	gca	ggt	gtg	gcc	atg 1245	acc	aac	ctg	cct	3803
Met	Gly	Ser	Ala	Val	Phe	Ala	Gly	Val	Ala	Met	Thr	Asn	Leu	Pro	
ggc 1250	atc	ctt	gtc	ctg	ggc 1255	ctc	gcc	aag	gcc	cag 1260	ctc	att	cag	atc	3848
Gly	Ile	Leu	Val	Leu	Gly	Leu	Ala	Lys	Ala	Gln	Leu	Ile	Gln	Ile	
ttc 1265	ttc	ttc	cgc	ctc	aac 1270	ctc	ctg	atc	act	ctg 1275	ctg	ggc	ctg	ctg	3893
Phe	Phe	Phe	Arg	Leu	Asn	Leu	Leu	Ile	Thr	Leu	Leu	Gly	Leu	Leu	
cat 1280	ggc	ttg	gtc	ttc	ctg 1285	ccc	gtc	atc	ctc	agc 1290	tac	gtg	ggg	cct	3938
His	Gly	Leu	Val	Phe	Leu	Pro	Val	Ile	Leu	Ser	Tyr	Val	Gly	Pro	
gac 1295	gtt	aac	ccg	gct	ctg 1300	gca	ctg	gag	cag	aag 1305	cgg	gct	gag	gag	3983
Asp	Val	Asn	Pro	Ala	Leu	Ala	Leu	Glu	Gln	Lys	Arg	Ala	Glu	Glu	
gcg 1310	gtg	gca	gca	gtc	atg 1315	gtg	gcc	tct	tgc	cca 1320	aat	cac	ccc	tcc	4028
Ala	Val	Ala	Ala	Val	Met	Val	Ala	Ser	Cys	Pro	Asn	His	Pro	Ser	
cga 1325	gtc	tcc	aca	gct	gac 1330	aac	atc	tat	gtc	aac 1335	cac	agc	ttt	gaa	4073
Arg	Val	Ser	Thr	Ala	Asp	Asn	Ile	Tyr	Val	Asn	His	Ser	Phe	Glu	
ggg 1340	tct	atc	aaa	ggg	gct 1345	ggg	gcc	atc	agc	aac 1350	ttc	ttg	ccc	aac	4118
Gly	Ser	Ile	Lys	Gly	Ala	Gly	Ala	Ile	Ser	Asn	Phe	Leu	Pro	Asn	

aat ggg cgg cag ttc tga tacagccaga ggccctgtct aggctctatg 4166
Asn Gly Arg Gln Phe
1355

gccctgaacc aaagggttat ggggatcttc cttgtgactg ccccttgaca cacgccctcc 4226
tcaaataccta ggggaggcca ttcccatgag actgcctgtc actggaggat ggcctgctct 4286
tgaggatatcc aggcagcacc actgatggct cctctgctcc catagtgggt ccccagtttc 4346
caagtcacct aggccttggg cagtgcctcc tcctgggcct gggctctggaa gttggcagga 4406
acagacacac tccatgtttg tcccacactc actcactttc ctaggagccc acttctcatc 4466
caacttttcc cttctcagtt cctctctcga aagtcttaat tctgtgtcag taagtcttta 4526
acacgtagca gtgtccctga gaacacagac aatgaccact accctgggtg tgatatcaca 4586
ggaggccaga gagaggcaaa ggctcaggcc aagagccaac gctgtgggag gccggtcggc 4646
agccactccc tccagggcgc acctgcaggt ctgccatcca cggccttttc tggcaagaga 4706
agggccccagg aaggatgctc tcataaggcc caggaaggat gctctcataa gcaccttggt 4766
catggattag cccctcctgg aaaatggtgt tgggtttggt ctccagctcc aatacttatt 4826
aaggctgttg ctgccagtca aggccacca ggagtctgaa ggctgggagc tcttggggct 4886
gggctggtcc tcccatcttc acctcgggcc tggatcccag gcctcaaacc agcccaacce 4946
gagcttttgg acagctctcc agaagcatga actgcagtgg agatgaagat cctggctctg 5006
tgctgtgcac ataggtgttt aataaacatt tgttggcaga aaaaaaaaaa aaaaaaaaaa 5066
aaaaaaaaaa aaaaaaaaaa aaaaaa 5092

<210> 44

<211> 1359

<212> PRT

<213> Homo sapiens

<400> 44

Met Ala Glu Ala Gly Leu Arg Gly Trp Leu Leu Trp Ala Leu Leu Leu
1 5 10 15

Arg Leu Ala Gln Ser Glu Pro Tyr Thr Thr Ile His Gln Pro Gly Tyr
20 25 30

Cys Ala Phe Tyr Asp Glu Cys Gly Lys Asn Pro Glu Leu Ser Gly Ser
35 40 45

Leu Met Thr Leu Ser Asn Val Ser Cys Leu Ser Asn Thr Pro Ala Arg
50 55 60

Lys Ile Thr Gly Asp His Leu Ile Leu Leu Gln Lys Ile Cys Pro Arg
65 70 75 80

Leu Tyr Thr Gly Pro Asn Thr Gln Ala Cys Cys Ser Ala Lys Gln Leu
85 90 95

Val Ser Leu Glu Ala Ser Leu Ser Ile Thr Lys Ala Leu Leu Thr Arg
100 105 110

Cys Pro Ala Cys Ser Asp Asn Phe Val Asn Leu His Cys His Asn Thr
115 120 125

Cys Ser Pro Asn Gln Ser Leu Phe Ile Asn Val Thr Arg Val Ala Gln
130 135 140

Leu Gly Ala Gly Gln Leu Pro Ala Val Val Ala Tyr Glu Ala Phe Tyr
145 150 155 160

Gln His Ser Phe Ala Glu Gln Ser Tyr Asp Ser Cys Ser Arg Val Arg
165 170 175

Val Pro Ala Ala Ala Thr Leu Ala Val Gly Thr Met Cys Gly Val Tyr
180 185 190

Gly Ser Ala Leu Cys Asn Ala Gln Arg Trp Leu Asn Phe Gln Gly Asp
195 200 205

Thr Gly Asn Gly Leu Ala Pro Leu Asp Ile Thr Phe His Leu Leu Glu
210 215 220

Pro Gly Gln Ala Val Gly Ser Gly Ile Gln Pro Leu Asn Glu Gly Val
225 230 235 240

Ala Arg Cys Asn Glu Ser Gln Gly Asp Asp Val Ala Thr Cys Ser Cys
245 250 255

Gln Asp Cys Ala Ala Ser Cys Pro Ala Ile Ala Arg Pro Gln Ala Leu
260 265 270

Asp Ser Thr Phe Tyr Leu Gly Gln Met Pro Gly Ser Leu Val Leu Ile

275					280					285					
Ile	Ile	Leu	Cys	Ser	Val	Phe	Ala	Val	Val	Thr	Ile	Leu	Leu	Val	Gly
	290					295					300				
Phe	Arg	Val	Ala	Pro	Ala	Arg	Asp	Lys	Ser	Lys	Met	Val	Asp	Pro	Lys
305					310					315					320
Lys	Gly	Thr	Ser	Leu	Ser	Asp	Lys	Leu	Ser	Phe	Ser	Thr	His	Thr	Leu
				325					330					335	
Leu	Gly	Gln	Phe	Phe	Gln	Gly	Trp	Gly	Thr	Trp	Val	Ala	Ser	Trp	Pro
			340					345					350		
Leu	Thr	Ile	Leu	Val	Leu	Ser	Val	Ile	Pro	Val	Val	Ala	Leu	Ala	Ala
		355					360					365			
Gly	Leu	Val	Phe	Thr	Glu	Leu	Thr	Thr	Asp	Pro	Val	Glu	Leu	Trp	Ser
	370					375					380				
Ala	Pro	Asn	Ser	Gln	Ala	Arg	Ser	Glu	Lys	Ala	Phe	His	Asp	Gln	His
385					390					395					400
Phe	Gly	Pro	Phe	Phe	Arg	Thr	Asn	Gln	Val	Ile	Leu	Thr	Ala	Pro	Asn
				405					410					415	
Arg	Ser	Ser	Tyr	Arg	Tyr	Asp	Ser	Leu	Leu	Leu	Gly	Pro	Lys	Asn	Phe
			420					425					430		
Ser	Gly	Ile	Leu	Asp	Leu	Asp	Leu	Leu	Leu	Glu	Leu	Leu	Glu	Leu	Gln
		435					440					445			
Glu	Arg	Leu	Arg	His	Leu	Gln	Val	Trp	Ser	Pro	Glu	Ala	Gln	Arg	Asn
	450					455					460				
Ile	Ser	Leu	Gln	Asp	Ile	Cys	Tyr	Ala	Pro	Leu	Asn	Pro	Asp	Asn	Thr
465					470					475					480
Ser	Leu	Tyr	Asp	Cys	Cys	Ile	Asn	Ser	Leu	Leu	Gln	Tyr	Phe	Gln	Asn
				485					490					495	
Asn	Arg	Thr	Leu	Leu	Leu	Leu	Thr	Ala	Asn	Gln	Thr	Leu	Met	Gly	Gln
			500					505					510		

Thr Ser Gln Val Asp Trp Lys Asp His Phe Leu Tyr Cys Ala Asn Ala
515 520 525

Pro Leu Thr Phe Lys Asp Gly Thr Ala Leu Ala Leu Ser Cys Met Ala
530 535 540

Asp Tyr Gly Ala Pro Val Phe Pro Phe Leu Ala Ile Gly Gly Tyr Lys
545 550 555 560

Gly Lys Asp Tyr Ser Glu Ala Glu Ala Leu Ile Met Thr Phe Ser Leu
565 570 575

Asn Asn Tyr Pro Ala Gly Asp Pro Arg Leu Ala Gln Ala Lys Leu Trp
580 585 590

Glu Glu Ala Phe Leu Glu Glu Met Arg Ala Phe Gln Arg Arg Met Ala
595 600 605

Gly Met Phe Gln Val Thr Phe Met Ala Glu Arg Ser Leu Glu Asp Glu
610 615 620

Ile Asn Arg Thr Thr Ala Glu Asp Leu Pro Ile Phe Ala Thr Ser Tyr
625 630 635 640

Ile Val Ile Phe Leu Tyr Ile Ser Leu Ala Leu Gly Ser Tyr Ser Ser
645 650 655

Trp Ser Arg Val Met Val Asp Ser Lys Ala Thr Leu Gly Leu Gly Gly
660 665 670

Val Ala Val Val Leu Gly Ala Val Met Ala Ala Met Gly Phe Phe Ser
675 680 685

Tyr Leu Gly Ile Arg Ser Ser Leu Val Ile Leu Gln Val Val Pro Phe
690 695 700

Leu Val Leu Ser Val Gly Ala Asp Asn Ile Phe Ile Phe Val Leu Glu
705 710 715 720

Tyr Gln Arg Leu Pro Arg Arg Pro Gly Glu Pro Arg Glu Val His Ile
725 730 735

Gly Arg Ala Leu Gly Arg Val Ala Pro Ser Met Leu Leu Cys Ser Leu
740 745 750

Ser Glu Ala Ile Cys Phe Phe Leu Gly Ala Leu Thr Pro Met Pro Ala
755 760 765

Val Arg Thr Phe Ala Leu Thr Ser Gly Leu Ala Val Ile Leu Asp Phe
770 775 780

Leu Leu Gln Met Ser Ala Phe Val Ala Leu Leu Ser Leu Asp Ser Lys
785 790 795 800

Arg Gln Glu Ala Ser Arg Leu Asp Val Cys Cys Cys Val Lys Pro Gln
805 810 815

Glu Leu Pro Pro Pro Gly Gln Gly Glu Gly Leu Leu Leu Gly Phe Phe
820 825 830

Gln Lys Ala Tyr Ala Pro Phe Leu Leu His Trp Ile Thr Arg Gly Val
835 840 845

Val Leu Leu Leu Phe Leu Ala Leu Phe Gly Val Ser Leu Tyr Ser Met
850 855 860

Cys His Ile Ser Val Gly Leu Asp Gln Glu Leu Ala Leu Pro Lys Asp
865 870 875 880

Ser Tyr Leu Leu Asp Tyr Phe Leu Phe Leu Asn Arg Tyr Phe Glu Val
885 890 895

Gly Ala Pro Val Tyr Phe Val Thr Thr Leu Gly Tyr Asn Phe Ser Ser
900 905 910

Glu Ala Gly Met Asn Ala Ile Cys Ser Ser Ala Gly Cys Asn Asn Phe
915 920 925

Ser Phe Thr Gln Lys Ile Gln Tyr Ala Thr Glu Phe Pro Glu Gln Ser
930 935 940

Tyr Leu Ala Ile Pro Ala Ser Ser Trp Val Asp Asp Phe Ile Asp Trp
945 950 955 960

Leu Thr Pro Ser Ser Cys Cys Arg Leu Tyr Ile Ser Gly Pro Asn Lys
965 970 975

Asp Lys Phe Cys Pro Ser Thr Val Asn Ser Leu Asn Cys Leu Lys Asn
980 985 990

Cys Met Ser Ile Thr Met Gly Ser Val Arg Pro Ser Val Glu Gln Phe
995 1000 1005

His Lys Tyr Leu Pro Trp Phe Leu Asn Asp Arg Pro Asn Ile Lys
1010 1015 1020

Cys Pro Lys Gly Gly Leu Ala Ala Tyr Ser Thr Ser Val Asn Leu
1025 1030 1035

Thr Ser Asp Gly Gln Val Leu Asp Thr Val Ala Ile Leu Ser Pro
1040 1045 1050

Arg Leu Glu Tyr Ser Gly Thr Ile Ser Ala His Cys Asn Leu Tyr
1055 1060 1065

Leu Leu Asp Ser Ala Ser Arg Phe Met Ala Tyr His Lys Pro Leu
1070 1075 1080

Lys Asn Ser Gln Asp Tyr Thr Glu Ala Leu Arg Ala Ala Arg Glu
1085 1090 1095

Leu Ala Ala Asn Ile Thr Ala Asp Leu Arg Lys Val Pro Gly Thr
1100 1105 1110

Asp Pro Ala Phe Glu Val Phe Pro Tyr Thr Ile Thr Asn Val Phe
1115 1120 1125

Tyr Glu Gln Tyr Leu Thr Ile Leu Pro Glu Gly Leu Phe Met Leu
1130 1135 1140

Ser Leu Cys Leu Val Pro Thr Phe Ala Val Ser Cys Leu Leu Leu
1145 1150 1155

Gly Leu Asp Leu Arg Ser Gly Leu Leu Asn Leu Leu Ser Ile Val
1160 1165 1170

Met Ile Leu Val Asp Thr Val Gly Phe Met Ala Leu Trp Gly Ile
1175 1180 1185

Ser Tyr Asn Ala Val Ser Leu Ile Asn Leu Val Ser Ala Val Gly
1190 1195 1200

Met Ser Val Glu Phe Val Ser His Ile Thr Arg Ser Phe Ala Ile

1205 1210 1215
Ser Thr Lys Pro Thr Trp Leu Glu Arg Ala Lys Glu Ala Thr Ile
1220 1225 1230
Ser Met Gly Ser Ala Val Phe Ala Gly Val Ala Met Thr Asn Leu
1235 1240 1245
Pro Gly Ile Leu Val Leu Gly Leu Ala Lys Ala Gln Leu Ile Gln
1250 1255 1260
Ile Phe Phe Phe Arg Leu Asn Leu Leu Ile Thr Leu Leu Gly Leu
1265 1270 1275
Leu His Gly Leu Val Phe Leu Pro Val Ile Leu Ser Tyr Val Gly
1280 1285 1290
Pro Asp Val Asn Pro Ala Leu Ala Leu Glu Gln Lys Arg Ala Glu
1295 1300 1305
Glu Ala Val Ala Ala Val Met Val Ala Ser Cys Pro Asn His Pro
1310 1315 1320
Ser Arg Val Ser Thr Ala Asp Asn Ile Tyr Val Asn His Ser Phe
1325 1330 1335
Glu Gly Ser Ile Lys Gly Ala Gly Ala Ile Ser Asn Phe Leu Pro
1340 1345 1350
Asn Asn Gly Arg Gln Phe
1355

<210> 45

<211> 4471

<212> DNA

<213> Mus musculus

<400> 45
ggatcacttc ctggctctgg gatggcagct gcctggcagg gatggctgct ctgggccctg 60
ctcctgaatt cggcccaggg tgagctctac acaccactc acaaagctgg cttctgcacc 120
ttttatgaag agtgtgggaa gaaccagag ctttctggag gcctcacatc actatccaat 180

atctcctgct tgtctaatac cccagccccg ccatgtcaca ggtgaccacc tggctcttct 240
ccagcgcgtc tgtccccgcc tatacaatgg cccaatgac acctatgcct gttgctctac 300
caagcagctg gtgtcattag acagtagcct gtctatcacc aaggccctcc ttacaogctg 360
cccggcatgc tctgaaaatt ttgtgagcat aactgtcat aatacctgca gccctgacca 420
gagcctcttc atcaatgtta ctgcgctggg tcagcgggac cctggacage ttcctgctgt 480
ggtggcctat gaggcctttt atcaacgcag ttttgcagag aaggcctatg agtcctgtag 540
ccgggtgcgc atccctgcag ctgcctcgct ggctgtgggc agcatgtgtg gagtgtatgg 600
ctctgccctc tgcaatgctc agcgcctggc tcaacttcca aggagacaca gggaatggcc 660
tggctccgct ggacatcacc ttccacctct tggagcctgg ccaggccctg gcagatggga 720
tgaagccact ggatgggaag atcaaaccct gcaatgagtc ccagggtgaa gactcggcag 780
cctgttctct ccaggactgt gcagcatcct gccctgtcat ccctccgcc ccggccctgc 840
gcccttcttt ctacatgggt cgaatgccag gctggctggc tctcatcctc atcttctactg 900
ctgtctttgt attgctctct gttgtccttg tgtatctccg agtggcttcc aacaggaaca 960
agaacaagac agcaggctcc caggaagccc ccaacctccc tcgtaagcgc agattctcac 1020
ctcacactgt ccttgggcgg ttcttcgaga gctggggaac aatgggtggc tcatggccac 1080
tcaactgtct ggcaactgtc ttcatagttg tgatagcctt gtcagtaggc ctgaccttta 1140
tagaactcac cacagaccct gtggaactgt ggtcggcccc taaaagccaa gcccggaag 1200
aaaaggcttt ccatgacgag cattttggcc ccttcttccg aaccaaccag atttttgtga 1260
cagctaagaa caggtccagc tacaagtacg actccctgct gctagggccc aagaacttca 1320
gtgggatcct atccctggac ttgctgcagg agctgttgga gctacaggag agacttcgac 1380
acctgcaagt gtggtcccat gaggcacagc gcaacatctc cctccaggac atctgctatg 1440
ctcccccaa accgcataac accagcctca ctgactgctg tgtcaacagc ctcttcaat 1500
acttcagaa caaccacaca ctctgctgc tcacagccaa ccagactctg aatggccaga 1560
cctccctggg ggactggaag gaccatttcc tctactgtgc caatgccctc ctacgtaca 1620
aagatggcac agccctggcc ctgagctgca tagctgacta cggggcgccgt gtcttccct 1680
tccttgctgt tgggggctac caagggacgg actactcgga ggcagaagcc ctgatcataa 1740
ccttctctat caataactac cccgctgatg atccccgcat ggcccacgcc aagctctggg 1800
aggaggcttt cttgaaggaa atgcaatcct tccagagaag cacagctgac aagttccaga 1860
ttgcgttctc agctgagcgt tctctggagg acgagatcaa tcgcactacc atccaggacc 1920
tgccctgtct tgccatcagc taccttatcg tcttctgtga catctccctg gccctgggca 1980

gctactccag atggagccga gttgcggtgg attccaaggc tactctgggc ctaggtgggg 2040
tggctgttgt gctgggagca gtcgtggctg ccatgggctt ctactcctac ctgggtgtcc 2100
cctcctctct ggtcatcatt caagtggtag ctttctctgt gctggctgtg ggagctgaca 2160
acatcttcat ctttgttctt gagtaccaga ggctgcctag gatgcccggg gagcagcgag 2220
aggtcacat tggcgcacc ctgggtagtg tggccccag catgctgctg tgcagcctct 2280
ctgaggccat ctgcttcttt ctagggggccc tgacctccat gccagctgtg aggacctttg 2340
ccttgacctc tggcttagca atcatctttg acttcctgct ccagatgaca gcctttgtgg 2400
ccctgctctc cctggatagc aagaggcagg aggctctctg ccccgacgtc gtgtgctgct 2460
tttcaagccg aaatctgccc ccaccgaaac aaaaagaagg cctcttactt tgcttcttcc 2520
gcaagatata cactcccttc ctgctgcaca gattcatccg ccctgttgtg ctgctgctct 2580
ttctggctct gtttggagca aacctctact taatgtgcaa catcagcgtg gggctggacc 2640
aggatctggc tctgccaag gattcctacc tgatagacta cttcctcttt ctgaaccggt 2700
acttggaagt ggggcctcca gtgtactttg acaccacctc aggctacaac ttttccaccg 2760
aggcaggcat gaacgccatt tgctctagtg caggctgtga gagcttctcc ctaaccaga 2820
aaatccagta tgccagtga ttcctaatc agtcttatgt ggctattgct gcctcctcct 2880
gggtagatga cttcatcgac tggctgacct catcctcctc ctgctgcccgc atttataccc 2940
gtggccccca taaagatgag ttctgtccct caacggatac ttccttcaac tgtctcaaaa 3000
actgcatgaa ccgcactctg ggtcccgtga gaccacaaac agaacagttt cataagtacc 3060
tgccctgggt cctgaatgat acgccaaca tcagatgtct taaagggggc ctagcagcgt 3120
atagaacctc tgtgaatttg atctcagatg gccagattat agcctcccag ttcattggcct 3180
accacaagcc cttacggaac tcacaggact ttacagaagc tctccgggca tcccggttgc 3240
tagcagccaa catcacagct gaactacgga aggtgcctgg gacagatccc aactttgagg 3300
tcttccctta cacgatctcc aatgtgttct accagcaata cctgacgggt ctccctgagg 3360
gaatcttcac tcttgcctc tgcttcgtgc ccacctttgt ggtctgctac ctectactgg 3420
gcctggacat acgctcagge atcctcaacc tgctctccat cattatgatc ctcgaggaca 3480
ccatcggcct catggctgtg tggggtatca gctacaatgc tgtgtccctc atcaaccttg 3540
tcacggcagt gggcatgtct gtggagttcg tgtcccacat taccgggtcc tttgctgtaa 3600
gcaccaagcc taccggctg gagagagcca aagatgctac tatcttcatg ggcagtgcgg 3660
tgtttgctgg agtggccatg accaacttcc cgggcacct catcctgggc tttgctcagg 3720

cccagcttat ccagattttc ttcttccgcc tcaacctcct gatcaccttg ctgggtctgc 3780
tacacggcct ggtcttcctg cccgttgtcc tcagctatct ggggccagat gttaaccaag 3840
ctctgggtact ggaggagaaa ctagccactg aggcagccat ggtctcagag ccttcttgcc 3900
cacagtaccc cttcccggt gatgcaaaca ccagtgacct atgttaacta aggctttaat 3960
ccagaattta tccctgaaat taatgctgct agcagctctc tgcccaaaag tgaccaaaag 4020
ttctaattga gtaggagctt gtccaggctc catgggtctt gctgataagg ggccacgagg 4080
gtcttccctc tgggtgtttc caaggcctgg ggaaagtgtg tccagaaaaa aattgctggc 4140
attcttgtcc tgaggcagcc agcactggcc actttgttgt cataggctcc cgaggccatg 4200
atcagattac ctctctgtgta aagagaatat cttgagtatt gtatgggatg tatcacatgt 4260
caattaataaa ggccatggcc tatggcttag gcaggaaata ggggtgtggaa catccaggag 4320
aagaaaggat tctgggataa aggacacttg ggaacgtgtg gcagtggtag ctgagcacag 4380
gtaattagcc atgtggcgaa atgtagatta atataaatgc atatctaagt tatgattcta 4440
gtctagctat atggccaagg tatttataaa t 4471

<210> 46

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 46

atgtaggtg agtctgaacc taccc

25

<210> 47

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 47
ggattgcatt tccttcaaga aagcc

25

<210> 48

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 48
tatggctctg ccctctgcaa tgctc

25

<210> 49

<211> 28

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 49
tcagcagcct ctgttccaca tacacttc

28

<210> 50

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 50
gttccacagg gtctgtggtg agttc

25

<210> 51

<211> 3996

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)..(3996)

<223> n is g or a or t or c

<400> 51

atggcngarg cnggnytnmg nggntggytn ytnntgggcny tnytnytnmg nytnngcncar	60
wsngarccnt ayacnacnat hcaycarccn ggntaytgyg cnttytayga ygartgyggn	120
aaraayccng arytnwsngg nwsnytnatg acnytnwsna aygtnwsntg yytnwsnaay	180
acnccngcnm gnaarathac nggngaycay ytnathytny tncaraarat htgyccnmgn	240
ytnstayacng gncnaayac ncargcntgy tgywsngcna arcarytngt nwsnytngar	300
gcnwsnytnw snathacnaa rgcnynytn acnmgtgyc cngcntgyws ngayaaytty	360
gtnaayytnc aytgycaaaa yacntgywsn ccnaaycarw snytnnttyat haaygtnacn	420
mgngtngcnc arytnngngc nggncarytn ccngcngtng tngcntayga rgcnttytay	480
carcaywsnt tygcngarca rwsntaygay wsntgywsnm gngtnmgngt nccngcngcn	540
gcnacnytn gngtnggnac natgtgyggn gntaygggnw sngcnyntg yaaygcncar	600
mgntggytna ayttycargg ngayacnggn aayggnytn cncnytn ga yathacntty	660
cayytnytn arcnggnca rgcngtnggn wsnggnathc arcnytnaa ygarggngtn	720
gcnmgntgya aygarwsnca rggngaygay gtngcnacnt gywsntgyca rgaytgygc	780
gcnwsntgy cngcnathgc nmgnccncar gcnynayw snacnttyta yytnngncar	840
atgccnggnw snytngtnyt nathathath ytnntgywsng tnttygcngt ngtnacnath	900
ytnytngtng gnttymngt ngcncngcn mgngayaarw snaaratggt ngayccnaar	960
aarggnacnw snytnwsnga yaarytnwsn ttywsnacnc ayacnytnyt nggncartty	1020
ttycarggnt ggggnacntg ggtngcnwsn tggcnytna cnathytngt nytnwsngtn	1080
athcngtng tngcnytnge ngcnggnytn gtnttyacng arytnacnac ngayccngtn	1140
garytnntggw sngcncnaa ywsncargcn mgnwsngara argcnttyca ygaycarcay	1200

ttyggncnt tyttymgnac naaycargtn athytnacng cncnaaymg nwsnwsntay 1260
mgntaygayw snytnytnyt nggncnaar aayttywsng gnathytnga yytngayytn 1320
ytntyngary tnytngaryt ncargarmgn ytnmgncayy tncargtntg gwsnccngar 1380
gncarmgna ayathwsnyt ncargayath tgytaygcnc cnytnaaycc ngayaayacn 1440
wsnytntayg aytgytgyat haaywsnytn ytncartayt tycaraayaa ymgnaacnytn 1500
ytntytnyna cngcnaayca racnytnatg ggncaracnw snargtnga ytggargay 1560
cayttyytn aytgygcnaa ygcncnytn acnttyaarg ayggnaacngc nytnngcnytn 1620
wsntgyatgg cngaytaygg ngcncngtn ttyccnttyy tngcnathgg nggntayaar 1680
ggnaargayt aywsngargc ngargcnytn athatgacnt tywsnytnaa yaaytayccn 1740
gcngngayc cnmgnytnge ncargcnaar ytnctggarg argcnttyyt ngargaratg 1800
mgngcnttyc armgnmgnt ggcnngnatg ttcargtna cnttyacngc ngarmgnwsn 1860
ytngargayg arathaaymg nacnacngcn gargayytn cnathttygc nacnwsntay 1920
athgtnatht tyytnatay hwsnytnge ytnngnwsnt aywsnwsntg gwsnmngtn 1980
atggtngayw snaargcnac nytnngnytn ggngngtnge cngtnngtnyt ngngngcngtn 2040
atggcngcna tgggnttytt ywsntayytn ggnathmgw snwsnytnge nathytnear 2100
gtngtnccnt tyytnngtnyt nwsngtnngn gcngayaaya thtttyatht ygtnytngar 2160
taycarmgny tncnmngmg nccngngar ccnmngarg tncayathgg nmngcnytn 2220
ggnmngtnge cncnwsnat gytnytnge wsnytnwsng argcnathtg ytttytytn 2280
ggngcnytna cncnatgcc ngcngtnmg acnttygcny tnacnwsngg nytnngcngtn 2340
athytngeyt tyytnytnca ratgwsngcn ttygtngcny tnytnwsnyt ngaywsnaar 2400
mgncargarg cnwsnmgny ngaygtntgy tgytgytna arccncarga rytncncn 2460
ccngncarg gngarggny nytnytnngn ttyttycara argentayge nccnttytn 2520
ytncaytgga thacnmngg ngtnngtnyt ytnytnntty tngcnytnnt yggngtnwsn 2580
ytnaywsna tgtgycaat hwsngtnngn ytngeyarg arytngeyt nccnargay 2640
wsntayytny tngaytaytt yytnnttytn aaymgntayt tygargtnge ngcncngtn 2700
tayttygtna cnacnytnge ntayaaytty wsnwsngarg cngnatgaa ygcnathtgy 2760
wsnwsngcng gntgyaayaa yttywsntty acncaraara thcartayge nacngartty 2820
ccngarcarw sntayytnge nathccngcn wsnwsntggg tngaygaytt yathgaytg 2880
ytnacncnw snwsntgyt ymgnytnay athwsnggnc cnaayaarga yaarttytg 2940

ccnwsnacng tnaaywsnyt naaytgyytn aaraaytgya tgwsnathac natgggnwsn	3000
gtnmgnccnw sngtngarca rttycayaar tayytnccnt ggattytnaa ygaymgncn	3060
aayathaart gyccnaargg nggnytnngcn gentaywsna cnwsngtnaa yytnacnwsn	3120
gayggncarg tnytnngcnws nmgnattyatg gentaycaya arccnytnaa raaywsncar	3180
gaytayacng argcnytnmg ngcngcnmgn garytnngcn cnaayathac ngcngayytn	3240
mgnaargtnc cnggnacnga yccngcntty gargnttyc cntayacnat hacnaaygtn	3300
ttytaygarc artayytnac nathytnccn garggnytn tyatgytnws nytntgyytn	3360
gtncnccnact tygcngtnws ntgyytnytn ytnngnytn ayytnmgcnws nggnytnytn	3420
aayytnytnw snathgtnat gathytnngtn gayacngtn gnttyatggc nytntgggay	3480
athwsntaya aygcngtnws nytnathaay ytnngtnwsng cngtnngnat gwsngtngar	3540
ttygtnwsnc ayathacnmg nwsnttygcn athwsnacna arccnccntg gytnngarmgn	3600
gcnaargarg cnacnathws natgggnwsn gcngnttyg cnggngtnge natgacnaay	3660
ytnccnggna thytnngtnyt nggnytnngcn aargncary tnathcarat httyttytty	3720
mgnytnaayy tnytnathac nytnytnngn ytnytncaay gnytnngntt yytnccngtn	3780
athytnwsnt aygtnngncc ngaygtnaay ccngcnytn cnytnngarca raarmngcn	3840
gargargcng tngcngcngt natggtnngcn wsntgyccna aycayccnws nmgnngtnwsn	3900
acngcngaya ayathtaygt naaycaywsn ttygarggnw snathaargg ngcnggngcn	3960
athwsnaayt tyytnccnaa yaayggnmgn cartty	3996

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 August 2005 (04.08.2005)

PCT

(10) International Publication Number
WO 2005/069900 A3

(51) International Patent Classification:

G01N 33/53 (2006.01) G01N 33/534 (2006.01)
G01N 33/532 (2006.01) C12Q 1/60 (2006.01)

(21) International Application Number:

PCT/US2005/001469

(22) International Filing Date: 14 January 2005 (14.01.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/537,341 16 January 2004 (16.01.2004) US

(71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, RY 60-30, Rahway, NJ 07065 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): GARCIA-CALVO, Margarita [US/US]; Westfield, NJ (US).

(74) Agent: FOLEY & LARDNER LLP; Washington Harbour, 3000 K. Street, N.W., Suite 500, Washington, DC 20007-5143 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the International search report:
8 June 2006

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF

(57) Abstract: The present invention provides human, rat and mouse NPC1L1 polypeptides and polynucleotides encoding the polypeptides. Methods for detecting ligands which bind to NPC1L1 and block intestinal cholesterol absorption are provided. Also included is a method of identifying ligands which bind to NPC1L1 using membranes derived from brush border membrane preparations. Compounds that bind to NPC1L1 can be used for inhibiting intestinal cholesterol absorption in a subject.

WO 2005/069900 A3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US05/01469

A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : G01N 33/53, 33/532, 33/534; C12Q 1/60 US CL : 435/7.1, 7.93, 11; 436/544, 545 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/7.1, 7.93, 11; 436/544, 545 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Continuation Sheet		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2004/0132058 A1 (Altmann et al.) 08 July 2004 (08.07.2004), especially paragraph [0011].	1, 7, 8, 10-13
A	US 2004/0161838 A1 (Altmann et al.) 19 August 2004 (19.08.2004), especially paragraph [0011].	1, 7, 8 and 10-13
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 19 January 2006 (19.01.2006)	Date of mailing of the international search report 13 MAR 2006	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer Shafiqul Haq Telephone No. 571-272-1600 <i>Janice Ford</i>	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US05/01469

Continuation of B. FIELDS SEARCHED Item 3:

APS, CAPLUS, STN search, East search.

Search terms : NPC3, NP1CIL1, azetidinone, ezetimibe, label, ligand, agonist, antagonist, detection, identification.